

**НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ
імені ІГОРЯ СІКОРСЬКОГО»**

*Факультет інформатики та обчислювальної техніки
Кафедра автоматизованих систем обробки інформації та управління*

"На правах рукопису"
УДК 004.031.43

До захисту допущено
В.о. завідувача кафедри

_____ Олександр ПАВЛОВ

“ _____ ” _____ 20 20 р.

МАГІСТЕРСЬКА ДИСЕРТАЦІЯ

на здобуття ступеня магістра

за освітньо-професійною програмою

«Інформаційні управляючі системи та технології»

зі спеціальності 126 «Інформаційні системи та технології»

на тему:

**«Інформаційно-пошукова система для підбору кредитних пропозицій
клієнтам торговельних мереж»**

Виконав:

студент VI курсу, групи ІС-92мп

Кас'янчук Андрій Сергійович _____

Керівник:

Доцент, к.т.н.,

Жураковська Оксана Сергіївна _____

Консультант:

професор, д.т.н., доцент,

Жаріков Едуард В'ячеславович _____

Рецензент:

Доцент каф. ТК, к.т.н., доцент

Лісовиченко Олег Іванович _____

Засвідчую, що у цій магістерській
дисертації немає запозичень з праць
інших авторів без відповідних посилань.

Студент _____

Київ – 2020 року

**НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ
імені ІГОРЯ СІКОРСЬКОГО»**

*Факультет інформатики та обчислювальної техніки
Кафедра автоматизованих систем обробки інформації та управління*

Рівень вищої освіти – *другий (магістерський)*

Спеціальність – *126 «Інформаційні системи та технології»*

Освітньо-професійна програма *«Інформаційні управляючі системи та технології»*

В.о.завідувача кафедри

_____ Олександр ПАВЛОВ

«__» _____ 2020 р.

**ЗАВДАННЯ
на магістерську дисертацію студенту**

Кас'янчук Андрій Сергійович

1. Тема дисертації «Інформаційно-пошукова система для підбору кредитних пропозицій клієнтам торговельних мереж», науковий керівник дисертації Жураковська Оксана Сергіївна, доцент, к.т.н., затверджені наказом по університету від «26» жовтня 2020 р. № 3132-с

2. Строк подання студентом дисертації “ 2 ” 12 20 20 р.

3. Об'єкт дослідження процес підбору кредитних пропозицій клієнтам торговельних мереж

4. Перелік завдань, які потрібно розробити:

- *провести аналіз і розглянути існуючі рішення задачі пошуку кредитних пропозицій;*
- *розробити алгоритм рішення задачі пошуку кредитних пропозицій;*
- *проаналізувати отримані результати;*
- *розробити програмну реалізацію.*

5. Орієнтовний перелік графічного (ілюстративного) матеріалу

Схема структурна бази даних 1, Схема структурна бази даних 2, Схема структурна бази даних 3, Схема структурна варіантів використання системи, Схема роботи Middle-ware системи, Схема роботи моніторингу сервісів банків, Схема роботи системи

6. Орієнтовний перелік публікацій

Дві публікації: дві тези доповіді на конференції

7. Консультанти розділів дисертації

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв

8. Дата видачі завдання “ 1 ” вересня 20 20 р.

Календарний план

№ з/п	Назва етапів виконання магістерської дисертації	Строк виконання етапів магістерської дисертації	Примітка
1	<i>Систематизація результатів огляду літератури</i>	<i>01.09 – 20.09</i>	
2	<i>Порівняльний аналіз існуючих методів розв’язання задачі</i>	<i>20.09 – 10.10</i>	
3	<i>Постановка та формалізація математичної моделі задачі</i>	<i>10.10 – 20.10</i>	
4	<i>Модифікація існуючих методів розв’язання задачі</i>	<i>20.10 – 30.10</i>	
5	<i>Розробка інформаційного та програмного забезпечення</i>	<i>01.11 – 10.11</i>	
7	<i>Проведення експериментальних досліджень розроблених алгоритмів</i>	<i>10.11 – 15.11</i>	
8	<i>Оформлення документації</i>	<i>15.11 – 20.11</i>	
9	<i>Подання роботи на попередній захист</i>	<i>20.11</i>	
10	<i>Подання роботи на основний захист</i>	<i>02.12</i>	

Студент

Андрій КАС’ЯНЧУК

Науковий керівник

Оксана ЖУРАКОВСЬКА

РЕФЕРАТ

Магістерська дисертація: 91 с., 10 рис., 27 табл., 1 додаток, 18 джерел.

Сьогодні на ринку банківських послуг дуже велика конкуренція, а вимоги клієнтів тільки ростуть. Споживачі, перед тим, як обирати кредитний продукт, хочуть порівняти між собою усі пропозиції на ринку та обрати найкращий серед запропонованих кредитних продуктів. Це зробити досить складно, потрібно відвідати кожен банк або переглянути відповідний веб-сайт банку, щоб зрозуміти, що він може запропонувати.

Особливо гостро ця проблема стоїть в сфері продажу товарів та послуг під споживчий кредит або так звану розстрочку. В Україні та у світі дуже мало торгових мереж, які можуть особисто видавати такі споживчі кредити. Загалом, магазини, в таких випадках, користуються послугами банків, проте для співпраці більше, ніж з одним банком потрібна налагоджена інфраструктура, і через це, з часом, обраний банк стає монополістом і починає завищувати процентні ставки на споживчий кредит. В такому випадку магазини починають втрачати потенційних клієнтів і гроші.

Залишається актуальною проблема зменшення часу на пошук та оформлення кредитних пропозицій. Процес прогнозування відповіді банку та процес підбору кредитних пропозицій значно зменшить час оформлення кредиту. Зменшення часу оформлення збільшить кількість клієнтів.

У даному дослідженні розглядається задача продажу кредитних продуктів у автоматичному режимі, без безпосередньої участі банку, використовуючи інформаційні технології. Створена інформаційна система дозволяє переглядати запропоновані кредитні продукти від усіх підключених до системи банків в одному місці. Система дозволяє підбирати кредитні продукти для клієнта і в онлайн режимі надсилати запити на отримання кредиту в усі банки по вибраним кредитним продуктам.

Зв'язок роботи з науковими програмами, планами, темами. Робота виконувалась на кафедрі автоматизованих систем обробки інформації та управління Національного технічного університету України «Київський політехнічний інститут ім. Ігоря Сікорського» в рамках теми «Математичні моделі та технології в СППР» (№ДР 0117U000914).

Мета дослідження: підвищення ефективності продажу товарів із залученням кредитних продуктів в торговельних мережах.

Об'єкт дослідження – процес підбору кредитних пропозицій клієнтам торговельних мереж.

Предмет дослідження – моделі і методи прогнозування відповіді банку та пошуку вигідних для клієнта кредитних пропозицій для купівлі товару в торговій мережі.

Методи дослідження. Методи дослідження, використані в роботі базуються на методах машинного навчання.

Наукова новизна одержаних результатів. Запропоновано схему побудови множини скорингових моделей, які використовуються для прогнозування результатів отримання клієнтом кредитних пропозицій від банків, а також покращення умов отримання кредитних пропозицій.

Публікації. Матеріали роботи опубліковані в журналі міжнародної конференції «PERSPECTIVES OF WORLD SCIENCE AND EDUCATION Osaka 2020» [10] та у V ВСЕУКРАЇНСЬКІЙ НАУКОВО-ПРАКТИЧНІЙ КОНФЕРЕНЦІЇ МОЛОДИХ ВЧЕНИХ ТА СТУДЕНТІВ «ІНФОРМАЦІЙНІ СИСТЕМИ ТА ТЕХНОЛОГІЇ УПРАВЛІННЯ» [11].

ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ, ІНФОРМАЦІЙНО-ПОШУКОВА СИСТЕМА, МОДЕЛЬ, ТОРГОВА МЕРЕЖА, СКОРИНГОВА МОДЕЛЬ, ПОШУК КРЕДИТНИХ ПРОПОЗИЦІЙ, ПІДТРИМКА ПРИЙНЯТТЯ РІШЕНЬ.

ABSTRACT

Master's dissertation consists 91 pages, 10 images, 27 tables, 1 appendix, 18 referring sources.

Today in the market of banking services there is a lot of competition, and customer requirements are only growing. Consumers, before choosing a credit product, want to compare all the offers on the market and choose the best among the offered credit products. This is quite difficult to do, you need to visit each bank or browse the relevant website of the bank to understand what it has to offer.

This problem is especially acute in the sale of goods and services on consumer credit or so-called installments. There are very few retail chains in Ukraine and in the world that can personally issue such consumer loans. In general, stores, in such cases, use the services of banks, but to cooperate with more than one bank requires a well-established infrastructure, and because of this, over time, the selected bank becomes a monopolist and begins to raise interest rates on consumer credit. In this case, stores begin to lose potential customers and money.

The problem of reducing the time for searching and processing loan proposals remains relevant. The process of forecasting the bank's response and the process of selecting loan proposals will significantly reduce the time of the loan. Reducing the clearance time will increase the number of customers.

This study examines the problem of selling credit products automatically, without the direct participation of the bank, using information technology. The created information system allows to view the offered credit products from all banks connected to the system in one place. The system allows you to select loan products for the client and send loan requests online to all banks for selected loan products.

Relationship with working with scientific programs, plans, topics. The work was performed at the Department of Automated Information Processing and Control Systems of the National Technical University of Ukraine "Kyiv Polytechnic Institute.

Igor Sikorsky” in the framework of the topic “Mathematical models and technologies in DSS” (DR 0117U000914).

The purpose of the study: to increase the efficiency of sales of goods involving credit products in retail chains.

The object of research is the process of selection of credit offers to customers of retail chains.

The subject of research - models and methods of forecasting the bank's response and finding favorable for the customer credit proposals for the purchase of goods in the retail network.

Research methods. The research methods used in the work are based on machine learning methods.

Scientific novelty of the obtained results. The scheme of construction of a set of scoring models which are used for forecasting of results of reception by the client of credit offers from banks, and also improvement of conditions of reception of credit offers is offered.

Publications. The materials of the work were published in the journal of the international conference "PERSPECTIVES OF WORLD SCIENCE AND EDUCATION Osaka 2020" [10] and in the V ALL-UKRAINIAN SCIENTIFIC-PRACTICAL CONFERENCE OF YOUNG SCIENTISTS AND STUDENTS [11].

INFORMATION TECHNOLOGIES, INFORMATION-SEARCH SYSTEM, MODEL, TRADE NETWORK, SCORING MODEL, SEARCH FOR CREDIT OFFERS, DECISION MAKING SUPPORT.

ЗМІСТ

ВСТУП.....	9
1 ОГЛЯД СУЧАСНИХ ДОСЛІДЖЕНЬ В ОБЛАСТІ ПОШУКУ КРЕДИТОРІВ ДЛЯ КЛІЄНТІВ ТОРГІВЕЛЬНИХ МЕРЕЖ.....	11
1.1 Огляд існуючих систем і підходів	11
1.1.1 <i>Поняття скорингу.....</i>	13
1.1.2 <i>Історія розвитку скорингу.....</i>	16
1.1.3 <i>Обмеження, пов'язані із застосуванням скорингу</i>	18
1.1.4 <i>Перспективи розвитку скорингу в Україні</i>	19
1.1.5 <i>Огляд рішень для банківського аналізу</i>	21
1.2 ОПИС БІЗНЕС ПРОЦЕСІВ	28
1.2.1 <i>Опис процесу діяльності.....</i>	28
1.2.2 <i>Актори і функції.....</i>	29
1.2.3 <i>Структура бізнес-процесів.....</i>	32
1.3 ОПИС ПОСТАНОВКИ ЗАДАЧІ	34
1.3.1 <i>Призначення розробки</i>	34
1.3.2 <i>Цілі та задачі розробки</i>	34
Висновок до розділу	35
2 РОЗРОБКА РІШЕНЬ ДЛЯ ПОШУКУ КРЕДИТОРІВ В ТОРГІВЕЛЬНИХ МЕРЕЖАХ	36
2.1 ЗМІСТОВНА ПОСТАНОВКА ЗАДАЧІ.....	36
2.2 МАТЕМАТИЧНА МОДЕЛЬ.....	36
2.3 ОГЛЯД МЕТОДІВ РОЗВ'ЯЗАННЯ	37
2.3.1 <i>Лінійна ймовірнісна модель.....</i>	38
2.3.2 <i>Логістична регресія.....</i>	39
2.4 РОЗРОБКА АЛГОРИТМУ РОЗВ'ЯЗАННЯ	41
2.4.1 <i>Пошук коефіцієнтів логістичної регресії</i>	41
2.4.2 <i>Скорингова карта</i>	43
2.4.3 <i>Значущі ознаки</i>	43
2.4.4 <i>Категоріальні ознаки.....</i>	45
2.5 РОЗРОБКА МЕТОДУ РОЗВ'ЯЗАННЯ ЗАДАЧІ.....	46
2.6 РЕЗУЛЬТАТИ ДОСЛІДЖЕНЬ ЕФЕКТИВНОСТІ ПОБУДОВАНОЇ СКОРИНГОВОЇ МОДЕЛІ В ПРОГНОЗУВАННІ.....	53
2.6.1 <i>Оцінка скорингових моделей</i>	53
2.6.2 <i>Аналіз результатів</i>	56
Висновок до розділу	60
3 ОПИС ПРОГРАМНОГО ТА ТЕХНІЧНОГО ЗАБЕЗПЕЧЕННЯ.....	61
3.1 ЗАСОБИ РОЗРОБКИ.....	61
3.2 ВИМОГИ ДО ТЕХНІЧНОГО ЗАБЕЗПЕЧЕННЯ.....	63
3.2.1 <i>Загальні вимоги</i>	63
Висновок до розділу	64
4 РОЗРОБЛЕННЯ СТАРТАП-ПРОЕКТУ.....	65

4.1	ОПИС ІДЕЇ ПРОЕКТУ	66
4.2	АНАЛІЗ РИНКУ	66
4.3	РОЗРОБЛЕННЯ РИНКОВОЇ СТРАТЕГІЇ.....	72
4.4	ПОБУДОВА МАРКЕТИНГОВОЇ ПРОГРАМИ.....	74
	Висновок до розділу	77
	Висновки	78
	Перелік посилань	80
	Додаток А Графічний матеріал	83
	СХЕМА СТРУКТУРНА БАЗИ ДАНИХ 1	84
	СХЕМА СТРУКТУРНА БАЗИ ДАНИХ 2	85
	СХЕМА СТРУКТУРНА БАЗИ ДАНИХ 3	86
	СХЕМА РОБОТИ СИСТЕМИ	87
	СХЕМА СТРУКТУРНА ВАРІАНТІВ ВИКОРИСТАННЯ.....	88
	СХЕМА РОБОТИ MIDDLE-WARE СИСТЕМИ	89
	СХЕМА РОБОТИ МОНІТОРИНГУ СЕРВІСІВ БАНКІВ	90

ВСТУП

Споживче кредитування - послуга, яка допомагала ринку роздрібною торгівлі набирати обертів. До останнього часу обсяги кредитів, виданих в торгових точках, тільки росли, а кількість галузей, де застосовується кредитування, розширювалося. Простота отримання позики в порівнянні з оформленням звичайного кредиту в банку приваблива для покупців. А використання спеціалізованого інструменту, робить цей процес ще більш простим і швидким за рахунок автоматизації процесу кредитування. Можливість вирішити питання тут і зараз збільшує шанси на те, що клієнт скористається кредитом, якщо його заявка буде схвалена. Чим більше банків-партнерів розгляне заявку, тим більше шансів, що кредит або розстрочка покупцеві будуть схвалені. Велика кількість банків-партнерів вигідно не тільки продавцеві, а й покупцеві, який, в свою чергу, може вибрати з різних пропозицій вигідніше для себе.

На ринку банківських послуг щороку конкуренція росте, а вимоги клієнтів ростуть. Споживачі перед тим як обирати кредитний продукт хочуть порівняти між собою усі пропозиції на ринку та обрати найкращий серед запропонованих кредитних продуктів. Без спеціального інструментарію це зробити досить складно, потрібно відвідати кожен банк або переглянути відповідний веб-сайт банку, щоб зрозуміти, що він може запропонувати.

У даному дослідженні розглядаються процес продажу кредитних продуктів з використанням спеціалізованого інструментарію у автоматичному режимі, без безпосередньої участі банківських працівників в процесі. Створено інформаційну систему, яка дозволяє переглядати запропоновані кредитні продукти від усіх підключених до системи банків в одному місці. Система не тільки підбирає кредитні продукти для клієнта, а і в онлайн режимі надсилає запити на отримання кредиту в усі банки по вибраним кредитним продуктам.

Перед відправкою в банк анкети, внутрішньою скоринговою системою прогнозується відповідь кожного з банку.

1 ОГЛЯД СУЧАСНИХ ДОСЛІДЖЕНЬ В ОБЛАСТІ ПОШУКУ КРЕДИТОРІВ ДЛЯ КЛІЄНТІВ ТОРГІВЕЛЬНИХ МЕРЕЖ

1.1 Огляд існуючих систем і підходів

Продаж кредитних продуктів в торгових мережах є одним із способів збільшити прибутки як і торгових мереж, так і банків. Існує декілька способів реалізації цієї ідеї.

До вирішення проблеми було виявлено такі рішення. Банк надає доступ торговим мережам до інтерфейсів банків і дозволяє продавцям в магазині безпосередньо в онлайн режимі посилати запит на отримання кредиту в банк [1]. Даний спосіб дозволяє пришвидшити процес отримання кредитів і дозволяє прийняти більшу кількість клієнтів, що значно збільшить прибутки торгової мережі і банку. Проте в умовах високої банківської конкуренції, клієнт хоче обирати не серед лише одного банку, клієнт хоче обрати найвигіднішу пропозицію від банку. Тому дана ідея в умовах ринкової економіки з досить великою конкуренцією далеко не ідеальна і потребує доопрацювання. Даний спосіб комунікації з банками використовує маса торгових мереж в багатьох країнах світу, і в часи коли цей метод вперше застосовувався це була революція в банківській сфері, проте на сьогодні цей метод вважається застарілим і все більше банків відходять від нього.

Було виявлено і альтернативне рішення проблеми. Описана подібна ідея як і в попередньому рішенні, проте там пропонується створити незалежну від банків і торгових мереж систему, що змогла б пропонувати кредитні продукти банків, як на біржі. Після реалізації цієї ідеї, автор теоретично доводить, що даний спосіб значно збільшує швидкість отримання позитивної відповіді від банку, так як в даній системі [2] запити в банки будуть відправлятися одночасно, а не так як описано в попередньому рішенні, де продавець опитує банки по черзі. Передбачувалось, що продавець на торговій точці лише заповнить анкету з

даними про клієнта і відправить ці дані в банки для отримання позитивної або негативної відповіді. Після отримання відповіді, клієнт сам в праві обрати кредитний продукт, який найбільш його зацікавив. Банк надсилає пакет документів, клієнт їх підписує. Даний метод був запропонований давно, але тоді через брак технологій він так і не був реалізований. Проте з сьогоденнішніми інформаційними технологіями та з доступністю комп'ютерів можна реалізувати систему, яка базується на попередніх принципах. По прогнозам, дана система зможе реалізувати систему «кредит за 5 хвилин» в торгових мережах, що збільшить кількість клієнтів в пікові дні продаж, такі як «Чорна П'ятниця», передноворічні та інші передсвяткові дні.

Описану інформаційну систему [2] можна використовувати для обробки і аналізу даних. Від аналізу даних можна отримати не лише дані по прибутку, а і аналізувати потреби клієнтів. За допомогою аналізу, наприклад, можна дізнатись, які товари клієнти найчастіше беруть в кредит, середню суму кредитів тощо, це допоможе підвищити обсяг продажів і краще розуміти потреби клієнтів, а банкам зрозуміти, які кредитні продукти клієнти обирають частіше інших.

В попередньому підході не описано можливість продавати кредитні продукти онлайн через сайт магазину, що стане реалізовано в розробленій системі. Клієнт заповнює анкету на отримання кредиту, оператор колл-центру з ним зв'язується, все перевіряє і потім надсилає запити на отримання кредиту. Якщо один із банків дав позитивну відповідь, то клієнта запрошують до торгової точки для підписання документів і отримання товару.

Наведені підходи і системи мають *недоліки*. Було запропоновано використовувати інформаційні системи, проте самі ж автори не володіли компетентними знаннями в інформаційних технологіях, так як вони володіли виключно економічними знаннями. В даному дослідженні предметна область буде описана з точки зору проектування та розробки інформаційної системи.

Було розглянуто деякі підходи реалізації продажу кредитних продуктів в торгових мережах. Обидва рішення для реалізації процесу залучають інформаційні технології, де пропонується створити систему описану вище.

Проблема залишається відкритою, бо не було описано підхід до реалізації архітектури інформаційної системи, підходи розв'язанні з точки зору проектування та розробки інформаційної системи. В дисертації проаналізовано предметну область та описано бізнес-процеси.

На основі аналізу існуючих рішень та виявлених проблем було сформульовано вимоги до інформаційної системи, розробці якої присвячено роботу.

В запропонованій схемі вирішення задачі підбору кредитних пропозицій для прогнозування відповідей банків використовується скорингова модель. На ринку банківський послуг дана модель використовується давно і має різні варіанти реалізації. Розглянемо деякі із них.

1.1.1 Поняття скорингу

Якість оцінки кредитного ризику тісно пов'язана з підвищення прибутковості кредитних операцій. Варто видавати кредит чи ні, які відсотки слід встановлювати, який ліміт кредитування? Відповіді на ці питання можна отримати після класифікації клієнта за групами ризику, після чого банк приймає рішення [4].

Сьогодні існує два основні методи оцінки ризику кредитування, що застосовуються як окремо, так і в поєднанні один з одним:

- висновок експертів або кредитних інспекторів;
- автоматизовані системи скорингу.

Скоринг - це математична або статистична модель, за допомогою якої на основі кредитної історії «минулих» клієнтів банк намагається визначити ймовірність, що конкретний потенційний позичальник поверне кредит без затримок.

У західній банківській системі, коли людина звертається за кредитом, банк може мати таку інформацію для аналізу [4]:

- анкета, яку заповнює позичальник;
- інформація на даного позичальника з кредитного бюро - організації, в якій зберігається кредитна історія всього дорослого населення країни;
- дані рухів по рахунках, якщо мова йде про діючого клієнта банку.

Кредитні аналітики оперують такими поняттями: «характеристики» клієнтів (в математичній термінології - змінні, чинники) і «ознаки» - значення, які приймає змінна. Якщо уявити собі анкету, яку заповнює клієнт, то характеристиками є питання анкети (вік, сімейний стан, професія), а ознаками - відповіді на ці питання[4].

У самому спрощеному вигляді скорингова модель представляє собою зважену суму певних характеристик. В результаті виходить інтегральний показник (*score*); чим він вищий, тим вище надійність клієнта, і банк може впорядкувати своїх клієнтів по мірі зростання кредитоспроможності.

Інтегральний показник кожного клієнта порівнюється з якимсь числовим порогом, або лінією розділу, яка, по суті, є лінією беззбитковості і розраховується з відношення, скільки в середньому потрібно клієнтів, які платять вчасно, для того, щоб компенсувати збитки від одного боржника. Клієнтам з інтегральним показником вище цієї лінії видається кредит, клієнтам з інтегральним показником нижче цієї лінії - ні [4].

Проте складність полягає у визначенні, які характеристики слід включати в модель і які вагові коефіцієнти повинні їм відповідати. До вирішення проблеми є кілька підходів, які будуть розглянуті в розділі 2.

Скоринг виділяє ті характеристики, які найбільш тісно пов'язані з ненадійністю або, навпаки, з надійністю клієнта. Модель не вказує, чи поверне даний позичальник кредит, але дозволяє визначити, що в минулому люди цього віку, цієї ж професії, з таким же рівнем освіти і з таким же числом утриманців кредит не повертали. Тому на основі моделі можна буде сформулювати рекомендації не давати кредит цій людині.

У цьому полягає дискримінаційний (не в статистичному, а в соціальному значенні цього слова) характер скорингу. Якщо людина за формальними ознаками близький до групи з поганою кредитною історією, то йому кредит не дадуть. Тому навіть при дуже високому ступені використання автоматизованих систем скорингу здійснюється суб'єктивне втручання в разі, коли кредитний інспектор володіє додатковою інформацією, яка доводить, що людина, класифікована як ненадійна, насправді «хороша», і навпаки.

Які ж характеристики є найбільш «цінними» для прогнозування кредитного ризику? У Великобританії найбільш часто використовуються такі характеристики:

- вік;
- кількість дітей / утриманців;
- професія;
- професія чоловіка (дружини);
- дохід;
- дохід дружини (чоловіка);
- район проживання;
- вартість житла;

- наявність телефону;
- скільки років живе за даною адресою;
- скільки років працює на цій роботі;
- скільки років є клієнтом даного банку;
- наявність кредитної картки / чекової книжки.

В інших країнах набір характеристик, які найбільш тісно пов'язані з ймовірністю дефолту - ймовірністю, що позичальник не поверне кредит або затримається з виплатою, буде відрізнятися в силу національних економічних і соціально-культурних особливостей. Чим більш однорідна популяція клієнтів, на якій розробляється модель, тим точніше прогнозування дефолту. Тому очевидно, що не можна автоматично перенести модель з однієї країни в іншу або з одного банку в інший. Навіть всередині одного банку існують різні моделі для різних груп клієнтів та різних видів кредиту [4].

1.1.2 Історія розвитку скорингу

Скоринг є методом класифікації популяції на різні групи, коли невідома характеристика, яка розділяє ці групи (поверне клієнт кредит чи ні), але відомі інші характеристики. У статистиці ідея класифікації популяції на групи була розроблена Фішером в 1936 р на прикладі рослин. У 1941 р Девід Дюран вперше застосував цю методику до класифікації кредитів на «погані» і «хороші». В цей час відбувались всім відомі події Другої світової війни, коли більшість кредитних аналітиків було мобілізовано на фронт, і фінансові структури зіткнулися з потребою в терміновій заміні цих фахівців. Банки змусили своїх аналітиків перед відходом написати правила, якими слід було керуватися при прийнятті рішення про видачу кредиту, щоб аналіз міг проводитися нефахівцями [4]. Це і був прообраз майбутніх експертних систем.

Масове використання скорингу почалося з поширенням кредитних карт. В банки кожен день зверталось за кредитними картами все більше і більше людей і банкам нічого не залишалося, як автоматизувати процес прийняття кредитних рішень. Однак дуже скоро було оцінено не лише швидкість обробки кредитних заявок, але і якість оцінки кредитних ризиків. За деякими дослідженнями, після впровадження скорингових систем кількість непогашених кредитів знизилася до 50%.

У 1974 р в США був прийнятий Закон про надання рівних можливостей на отримання кредиту, який забороняв відмовляти у видачі кредиту на підставі таких показників: раса, колір шкіри, національне походження, вік, стать, сімейний стан, релігія, отримання соціальної допомоги, відстоювання прав споживачів. У Великобританії законодавство допускає використання інформації про вік і сімейний стан, але зате забороняє брати до уваги будь-які фізичні недоліки і інвалідність. Для кредитних організацій використання скорингових систем стало доказом виконання цих антидискримінаційних законів - у комп'ютера немає упереджень.

Крім встановлення принципів рівноправності в області кредитування, кредитне законодавство США, як і Закон про споживчий кредит, прийнятий у Великобританії в тому ж 1974 г., мали важливе значення для формування служби кредитних бюро. В таких бюро записується кредитна історія всіх людей, коли-небудь зверталися за позикою в будь-яку кредитну організацію країни.

У кредитних бюро зберігаються такі види даних:

- соціально-демографічні характеристики;
- судові рішення;
- індивідуальні дані позичальників, отримані від кредитних організацій за принципом «ти - мені, я - тобі», банк може отримувати інформацію про клієнтів інших банків, тільки якщо сам поставляє аналогічну інформацію.

Обсяг і характер інформації, що містяться в бюро, суворо регулюється законодавством кожної країни.

Значення кредитних бюро дуже велике, їх існування дозволяє банкам надавати кредити новим для банку клієнтам [4].

В даний час скоринг стає все більш популярним не тільки при оцінці ризику при різних видах кредиту, але і в інших областях: у маркетингу (для визначення ймовірності, що саме ця група клієнтів буде користуватися цим видом продукції), при роботі з боржниками (якщо клієнт затримується з черговим платежем, який метод впливу буде найбільш ефективним), при виявленні шахрайства з кредитними картками, при визначенні ймовірності, що клієнт може перебігти до конкурента і т. п.

1.1.3 Обмеження, пов'язані із застосуванням скорингу

У скорингу існує дві основні проблеми. Перша полягає в тому, що класифікація вибірки проводиться тільки на клієнтах, яким дали кредит. Ми ніколи не дізнаємося, як би повели себе клієнти, яким в кредиті було відмовлено: цілком можливо, що якась частина виявилася б цілком прийнятними позичальниками.

Але, як правило, відмова в кредиті проводиться на підставі досить серйозних причин. Банки фіксують ці причини відмови і зберігають інформацію про «відмовників». Це дозволяє їм відновлювати первинну популяцію клієнтів, які зверталися за кредитом.

Друга проблема полягає в тому, що люди з плином часу змінюються, змінюються і соціально-економічні умови, що впливають на поведінку людей. Тому скорингові моделі необхідно розробляти на вибірці з найбільш «свіжих» клієнтів, періодично перевіряти якість роботи системи і, коли якість погіршується, розробляти нову модель. На Заході нова модель розробляється в

середньому раз в півтора року, період між заміною моделі може варіюватися в залежності від того, наскільки стабільною була економіка в цей час. Для України, ймовірно, максимальним періодом буде півроку, та й то за умови, що в цей період не відбудеться ніяких кардинальних потрясінь типу подій 2014 р [4]. В даний час ведуться дослідження того, як вводити соціально-економічні характеристики в модель з тим, щоб вона служила довше.

1.1.4 Перспективи розвитку скорингу в Україні

В Україні використання скоринг-систем гальмується, насамперед, низькими обсягами кредитування. Але з економічним зростанням ситуація почне змінюватися.

Само по собі невелике в порівнянні з західними кредитними організаціями кількість позичальників перешкодою не є, необхідно лише стежити за кількістю характеристик по відношенню до величини вибірки.

Відсутність кредитних бюро, безумовно, також не сприяє розвитку скорингу. Але, з іншого боку, на Заході існує проблема перевірки достовірності інформації, яку людина вказує про себе в анкеті. В Україні велика частина такої інформації міститься в паспорті. Банкам досить мати паспортні дані та дані трудової книжки – це вихідний матеріал для аналізу.

Ще один несприятливий фактор - недостатня поширеність таких універсальних статистичних пакетів, як SAS і SPSS. Крім того, існують і інші програми, доступні за ціною, які можуть робити лінійну багатофакторну регресію, а для початку цього цілком достатньо.

Цілком ймовірно, що в Україні скоринг спочатку буде застосовуватися не для фізичних осіб, а для юридичних тому, що у банків накопичено набагато більше інформації про підприємства, при цьому використовуються бальні системи оцінки ризику різної складності і з різним рівнем автоматизації.

Відмінність бальної системи від скорингової полягає в тому, що в першій значимість того чи іншого коефіцієнта або фінансового показника визначається суб'єктивно, а в другій виробляється прив'язка коефіцієнтів до рівня ризику.

На Заході при кредитуванні юридичних осіб скоринг-моделі поширені не настільки широко, як в споживчому кредиті. Це пов'язано з тим, що для розробки моделі дуже важко набрати достатню кількість компаній, подібних один з одним: компанії сильно відрізняються за розміром, обороту, секторам економіки. Чим більше підприємство, тим важче підібрати аналогічні підприємства для порівняння.

В останні роки великі зрушення відбулися в розробці скоринг-моделей для малого бізнесу. Застосування скорингу для малого і середнього бізнесу виявилось можливим саме в силу великої кількості подібних між собою підприємств.

В Україні впровадження скорингу гальмується не так об'єктивними, скільки суб'єктивними причинами, пов'язаними з недовірливим ставленням банківських менеджерів до математичних і статистичних методів. Не так вже й багато потрібно, щоб почати аналізувати своїх клієнтів - кредитна історія минулих клієнтів і статистичний пакет, - а віддача буде колосальною. Західні банкіри виділяють такі переваги скорингових систем, по перше - зниження рівня неповернення кредиту. Далі зазначається швидкодія і неупередженість у прийнятті рішень, відсутність необхідності довготривалого навчання працівників банку.

Впровадження скорингу в Україні повинно здійснюватися поступово. Для початку можна зробити автоматизовану систему, яка б оцінювала позичальників та буде в автоматичному режимі відсівати клієнтів з явно «поганими» ризиками, а на аналіз кредитного комітету надавати «хороші» ризики. Не вводячи автоматизацію, можна знайти зв'язок окремих ознак клієнта з ймовірністю

невиплати. Наявність таких ознак може послужити суттєвою підтримкою кредитним інспекторам [4].

Отже, скоринг являє собою систему, яка оцінює кредитний ризик, що працює в автоматичному режимі, які широко використовуються в США і Західній Європі. В якості вихідного матеріалу для скорингу використовується інформація про клієнтів, які раніше користувались послугами банків, на основі якої за допомогою різних статистичних та нестатистичних методів класифікації робиться прогноз про платоспроможність майбутніх позичальників. Скоринг-системи дозволяють банківським працівникам швидко приймати рішення про кредитування, регулювати обсяги кредитування в залежності від ситуації на ринку і визначати оптимальне співвідношення між прибутковістю кредитних операцій і рівнем ризику.

1.1.5 Огляд рішень для банківського аналізу

Огляд системи SAS Enterprise Miner

Лідером аналізу даних в банківській сфері є програмне забезпечення SAS Enterprise Miner від компанії SAS Institute. Інтерфейс програмного забезпечення зображений на рисунку 1.1.



Рисунок 1.1 – Інтерфейс системи SAS Enterprise Miner

SAS Enterprise Miner - це рішення для створення точних прогнозних та описових моделей для великих обсягів даних з різних джерел організації. SAS є лідером у галузі програмного забезпечення та послуг для бізнес-аналітики та найбільшим незалежним постачальником на ринку бізнес-аналітики [19].

SAS Enterprise Miner застосовується для вирішення таких завдань, як виявлення випадків шахрайства, мінімізація фінансових ризиків, оцінка і прогнозування потреб в ресурсах, підвищення ефективності маркетингових кампаній і зниження відтоку клієнтів.

Гнучка система методів пристосована до вирішення завдань різної складності. Рішення перетворює первинні дані в інформацію для подальшого використання аналітиками, статистами, бізнес-менеджерами і IT-фахівцями. Цілісність процесу і використовуваних методів дозволяють різним фахівцям об'єднувати зусилля і підвищувати ефективність застосовуваних рішень [19].

SAS Enterprise Miner надає аналітикам і data scientists (фахівці з інтелектуального аналізу даних) автодокументоване проектне середовище, яке

прискорює швидкість розробки моделей і об'єднує всі етапи аналізу даних, допомагаючи отримати найкращі результати [19].

Рішення автоматизує тривалий процес скорингу і генерує код для всіх стадій впровадження моделі на мовах програмування SAS, C, Java або PMML. Такий код може бути в подальшому використаний безліччю інтерактивних і пакетних середовищ як всередині SAS, так і в веб-додатках, в базах даних і безпосередньо в бізнес-процесах.

За допомогою інструменту SAS Rapid Predictive Modeler бізнес-користувачі, що володіють лише початковими знаннями в моделюванні, генерують прогнозні моделі для різних бізнес-цілей [19]. Аналітичні результати можуть бути легко інтерпретовані на основі простих і зрозумілих графіків і таблиць, розкриваючи необхідну для прийняття рішень інформацію.

Метрики оцінки якості моделей, побудовані різними способами, можуть бути виведені в зведеній таблиці, що значно полегшує їх порівняння. Підсумкові діаграми процесу моделювання можуть використовуватися в якості шаблону, який зручно редагувати, оновлювати і застосовувати до нових бізнес-завдань, а опис моделі містить інформацію про те, який внесок внесла кожна незалежна змінна в підсумковий результат. Крім того, точність моделі ґрунтується на базі сучасних алгоритмів, що враховують галузеву специфіку методів, що гарантує високу ступінь стабільності і надійності результатів.

Переваги SAS Enterprise Miner [19]:

- просунуті методи прогнозного і описового моделювання;
- зручний і зрозумілий інтерфейс, що дозволяє користувачам самостійно створювати моделі аналізу і прогнозування;
- автоматизований процес регламентного застосування моделей;
- можливість пакетної обробки складних процесів;
- швидкий збір та підготовка даних, їх агрегація і дослідження;

- простота масштабування і налаштування рішення;
- висока продуктивність системи навіть при роботі з великим масивом розрізнених даних.

Огляд системи IBM SPSS

Продукт IBM SPSS Statistics (раніше відомий просто як «SPSS») надає широкі можливості для аналізу даних. Інтуїтивно зрозумілий інтерфейс програмного забезпечення включає в себе всі функції управління даними, статистичні процедури і засоби створення звітів для проведення аналізу будь-якого ступеня складності. Інтерфейс програмного застосунку зображений на рисунку 1.2.

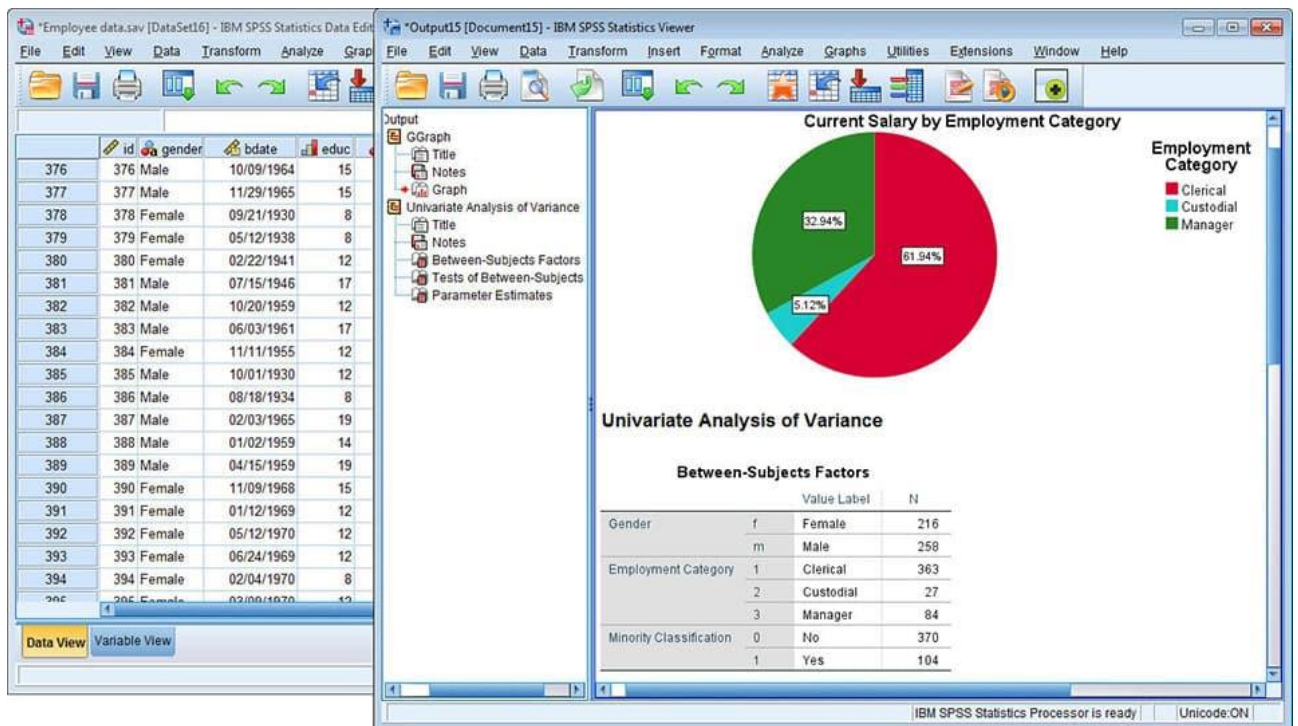


Рисунок 1.2 – Інтерфейс системи IBM SPSS

IBM SPSS Statistics і продукти IBM SPSS Amos, Data Collection, Collaboration і Deployment Services утворюють модульний, повністю інтегрований програмний комплекс, що охоплює всі етапи аналітичного процесу:

- планування дослідження;
- збір даних;
- доступ і керування даними;
- всебічний аналіз (від базових процедур виведення підсумків і класичної статистики до моделювання із застосуванням новітніх алгоритмів);
- створення звітів;
- зберігання та поширення результатів.

Інтеграція продуктів IBM SPSS в єдину лінійку дозволяє впевнено працювати, не стикаючись з проблемами переходу від одного програмного продукту до іншого.

Великий вибір процедур в базовому модулі IBM SPSS Statistics дає широкі можливості аналізу даних різних типів. Вбудовувані додаткові модулі розширюють аналітичні можливості настільки, наскільки це необхідно.

Фахівці в галузі аналізу даних відзначають логічність, продуманість і взаємопов'язаність компонентів програмного забезпечення IBM SPSS.

Огляд системи Deductor

Deductor є аналітичною платформою - основою для створення прикладних рішень в області аналізу даних. Реалізовані в Deductor технології дозволяють на базі єдиної архітектури пройти всі етапи побудови аналітичної системи: від

створення сховища даних до автоматичного підбору моделей і візуалізації отриманих результатів. Deductor складається з *шести* частин [10].

Deductor Warehouse багатовимірне сховище даних, що акумулює всю необхідну для аналізу предметної області інформацію. Використання єдиного сховища дозволяє забезпечити несуперечність даних і централізоване зберігання, а також автоматично забезпечує всю необхідну підтримку процесу аналізу. Deductor Warehouse оптимізований для вирішення саме аналітичних задач, що позитивно позначається на швидкості доступу до даних. У ряді випадків має сенс відмовитися від традиційного сховища і скористатися альтернативою - віртуальним сховищем Virtual Warehouse.

Deductor Studio програма, яка реалізує функції імпорту, обробки, візуалізації і експорту даних. Вона може функціонувати і без сховища, отримуючи інформацію з будь-яких інших джерел, але найкращим є їх спільне використання. В Deductor Studio включений повний набір механізмів, що дозволяє отримати інформацію з довільного джерела даних, провести весь цикл обробки (очищення, трансформацію даних, побудову моделей), відобразити отримані результати найбільш зручним чином (OLAP, таблиці, діаграми, дерева і т. Д.) і експортувати їх в найбільш поширені формати.

Deductor Viewer програма, орієнтована на кінцевого користувача і призначена для перегляду підготовлених за допомогою Deductor Studio звітів. Deductor Viewer дозволяє мінімізувати вимоги до користувача системи. Всі необхідні операції виконуються автоматично за допомогою підготовлених раніше сценаріїв обробки. Користувачеві Deductor Viewer потрібно тільки вибрати і налаштувати варіант відображення отриманих результатів.

Deductor Analytic Server служба, яка забезпечує віддалену аналітичну обробку даних. Вона дозволяє автоматично обробляти дані і перенавчати моделі на сервері, оптимізує виконання сценаріїв за рахунок кешування проектів і використання многопоточної обробки;

Deductor Integration Server служба, яка забезпечує інтеграційна взаємодія компонентів.

Deductor Client клієнт доступу до Deductor Server. Він забезпечує доступ до сервера з сторонніх додатків і управління його роботою. Реалізована в Deductor архітектура дозволяє домогтися максимальної гнучкості при створенні закінченого рішення. Завдяки цій архітектурі можна зібрати в одному аналітичному додатку всі необхідні інструменти аналізу і реалізувати автоматичне виконання підготовленого сценарію. Технологічна платформа включає кошти, що дозволяють максимально скоротити терміни розробки, швидко створювати і виводити на ринок нові прикладні рішення, а також адаптувати їх відповідно до вимог, що змінюються підприємств. Перші результати користувач отримує практично відразу, але при цьому можна легко нарощувати потужність рішення [10].

Кредитний брокер Фокстрот

На ринку України існує рішення для пошуку кредитних пропозицій від компанії Фокстрот. Програмний застосунок виступає, як посередник між банками та клієнтами торговельних мереж, що потребують отримання кредиту для купівлю товару. Кредитний брокер відсилає запити в підключені банки та знаходить пропозиції для клієнта, а також бере на себе ведення переговорів з кредиторами.

Кредитний брокер Фокстрот не використовує аналізу даних для прогнозування відповідей банків, тому є необхідність розробки системи, яка б будувала прогножуючу скорингову модель для передбачення відповідей банків.

Розглянуті рішення в повній мірі не реалізують функціонал, потрібний для вирішення задачі пошуку і підбору кредитного продукту.

1.2 Опис бізнес процесів

1.2.1 Опис процесу діяльності

Підхід передбачає, що клієнт особисто або за допомогою продавця на торговій точці, заповнить анкету з необхідними даними для банків. Анкету передбачається поділити на: основні дані, контакти, фінанси, робота. Після заповнення анкети, внутрішня скорингова система, на базі попередніх відповідей, аналізує анкету і надає клієнту прогнози на те, отримає клієнт кредит чи ні. Після аналізу, якщо система поверне негативну відповідь, буде виведено на екран поля в анкеті, після зміни яких, можливо буде отримати позитивну відповідь.

Внутрішня скорингова система для кожного банку будується на основі множини заявок і відповідей банків на ці заявки.

Після отримання прогнозу від внутрішньої скорингової системи, анкета відправляється в банк для аналізу банківською скоринговою системою. Протягом декількох хвилин банк повертає в систему результат роботи зовнішньої скорингової системи. Схема роботи зображена на рисунку 1.3.

У відповіді очікується результат (так або ні) та при позитивній відповіді множина кредитних продуктів.

Після отримання результатів, клієнту потрібно обрати банк, серед тих, що дали позитивну відповідь та обрати кредитний продукт. По кожному кредитному продукту буде представлена повна аналітика та буде рекомендовано найвигідніший продукт.

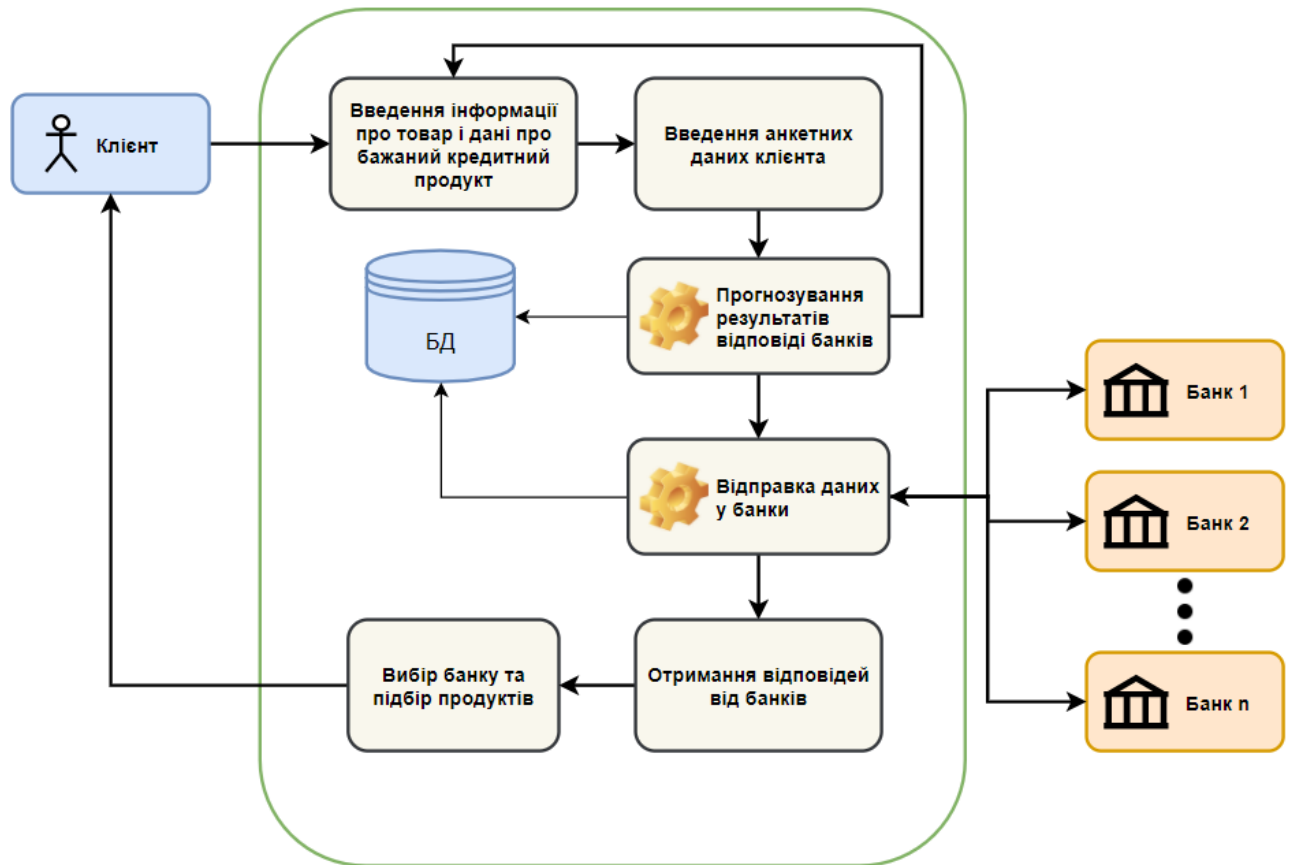


Рисунок 1.3 – Схема роботи системи

Вибравши кредитний продукт, система надсилає в банк відповідь, після чого система банку генерує пакет документів та надсилає в внутрішню систему. Клієнт підписує угоду, після чого в систему банку надсилається повідомлення про фіналізацію угоди.

1.2.2 Актори і функції

Функціональна модель системи базується на виконанні таких процесів:

- ведення товарів в системі;
- ведення кредитних продуктів в системі;
- ведення заявок в системі;
- ведення користувачів в системі;

- процес прогнозування відповіді банків;
- процес побудови актуальної скорингової моделі;
- процес відправки заявки в банки;
- процес обробки відповідей від банків.

Однією з основних задач є реалізація побудови скорингової моделі банків на основі отриманих раніше даних та прогнозування відповідей банків.

Система має таких акторів:

- продавець;
- банк (як система).

Клієнт безпосередньо в системі не є актором, а всю інформацію про клієнта вводить продавець.

Схема структурна варіантів використання наведена на рисунку 1.4.

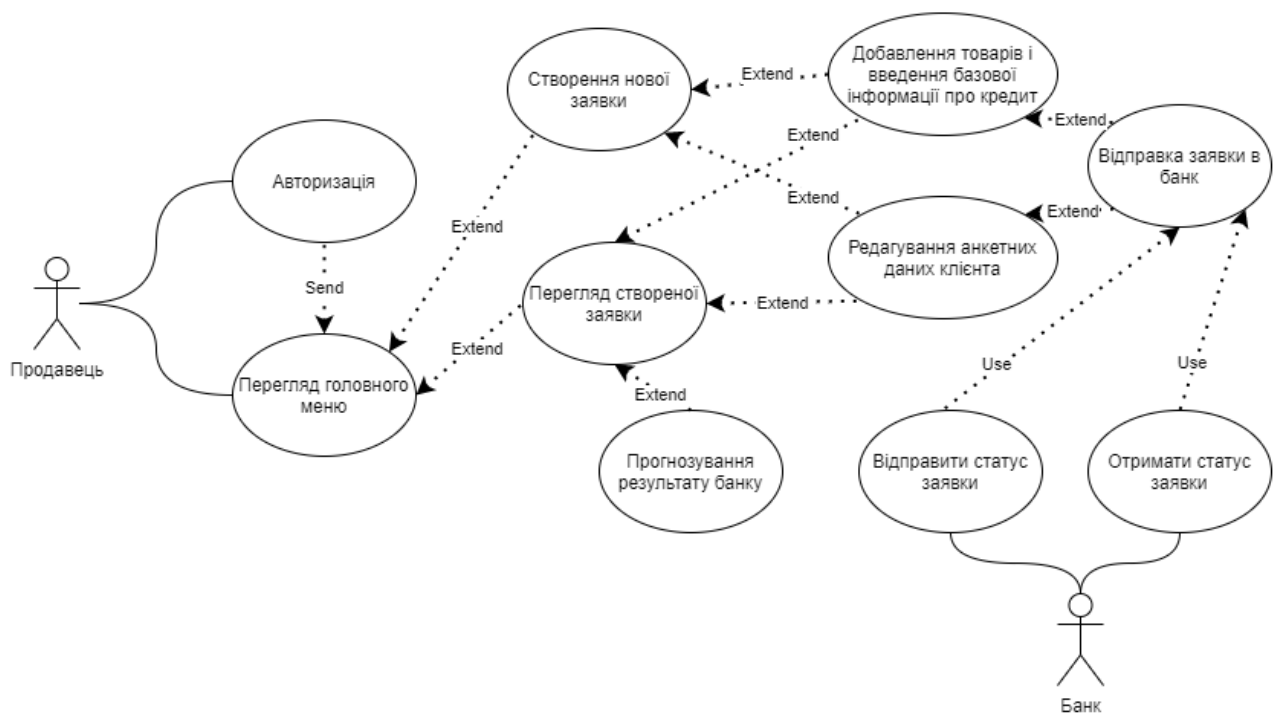


Рисунок 1.4 – Схема структурна варіантів використання системи

Таблиця 1.1 – Опис варіантів використання

Актор	Назва варіанту	Опис	Пріоритет
Продавець	Авторизація	Користувач авторизується для доступу до функціоналу системи.	Високий
	Перегляд головного меню	Користувач має можливість користуватись головним меню	Високий
	Створення нової заявки	Користувач має можливість створити нову заявку для отримання кредиту	Високий
	Перегляд створеної заявки	Користувач має можливість переглядати заявки зі своєї торгової точки	Високий
	Додавання товарів і введення базової інформації про кредит	Користувач має можливість додавати в заявку товари, які хоче купити клієнт та базові дані про кредит	Високий
	Редагування анкетних даних клієнта	Користувач має можливість додавати і редагувати анкетні дані клієнта	Високий
	Відправка заявки в банк	Користувач має можливість відправляти заявку в банк	Високий
Банк (як системи)	Відправити статус заявки	Банк може відправити статус заявки	Високий
	Отримати статус заявки	Банк може отримати статус заявки	Високий

1.2.3 Структура бізнес-процесів

Процес відправки заявки взаємодіє між двома акторами. Процес відправки заявки описаний в діаграмі діяльності на рисунку 1.5.

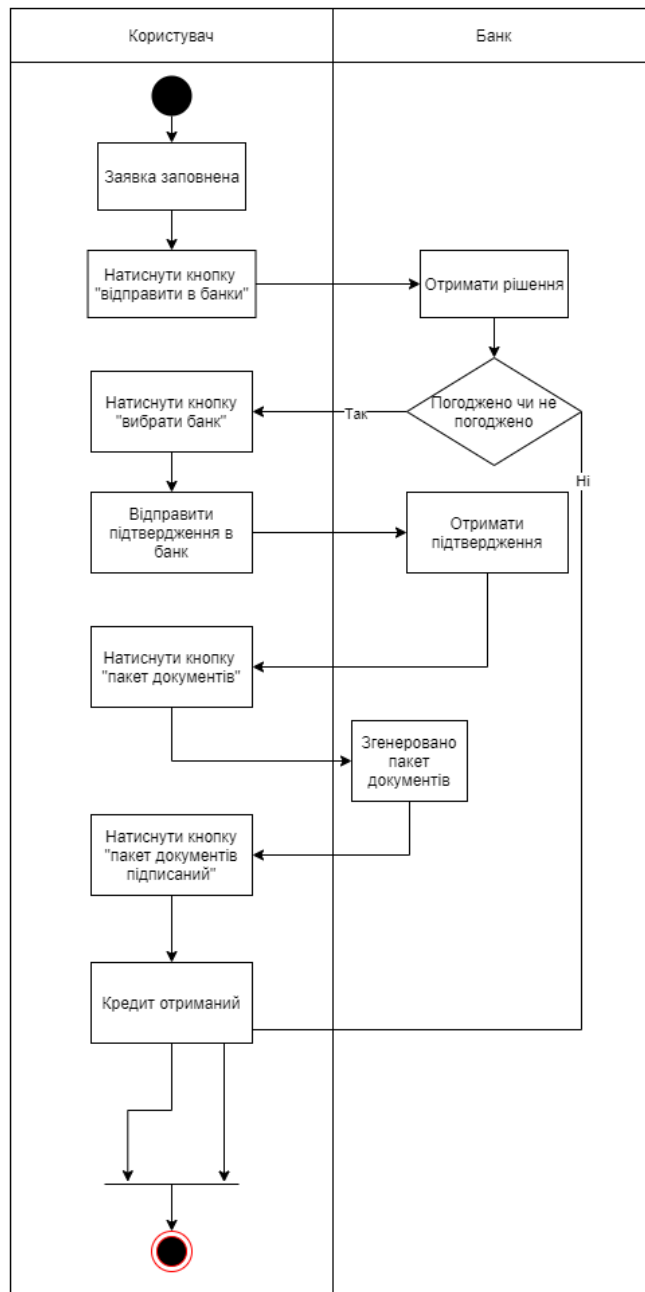


Рисунок 1.5 – Схема структурна діяльності процесу відправки заявки в банк

Таблиця 1.2 – Залежність функціональних вимог від кроків бізнес процесів

№	Крок бізнес-процесу	Функціональна вимога
1.	Авторизація користувача	1. Сервіс надає можливість користувачу авторизуватись 1.1. Сервіс надає можливість відправити запит на зміну паролю
2.	Створення заявки	2. Сервіс надає можливість створювати заявку на отримання кредиту 2.1. Можна додати в заявку товари 2.2. Можна задати перший внесок і термін бажаного кредиту 2.3. Можна редагувати поля анкети 2.4. Сервіс валідує введені дані
3.	Прогнозування	3. Сервіс надає можливість прогнозувати результати відповідей банків 3.1. Можна оцінити скоринг бал для кожного банку 3.2. Можна оцінити ймовірність отримання кредиту від кожного банку
4.	Система відправки заявки в банк	4. Сервіс надає можливість відправляти заявку в банк 4.1. Можна відмовитись від пропозицій 4.2. Можна вибрати банк і пропозицію 4.3. Можна запросити документи від банку 4.4. Можна фіналізувати згоду

1.3 Опис постановки задачі

1.3.1 Призначення розробки

Система призначена для пошуку і підбору кредитних пропозицій на основі централізованої взаємодії торгових мереж з банками.

1.3.2 Цілі та задачі розробки

Метою розробки є підвищення ефективності продажу товарів із залученням кредитних продуктів.

Для досягнення поставленої мети необхідно виконати наступні завдання:

- провести аналіз і розглянути існуючі рішення задачі пошуку кредитних пропозицій;
- провести дослідження для проектування системи пошуку кредитних пропозицій;
- розробити алгоритм побудови скорингової моделі для кожного банку;
- розробити алгоритм прогнозування відповіді банку;
- спроектувати базу даних для збереження анкет клієнтів та відповіді банків;
- розробити інтерфейс для взаємодії продавця і системи;
- виконати аналіз отриманих результатів;
- розробити програмну реалізацію розроблених алгоритмів.

Висновок до розділу

В даному розділі проаналізовано предметну область. Розглянуто поняття скорингу, історію розвитку скорингу, обмеження, які діють на скоринг та перспективи розвитку в Україні.

Проведено аналіз і розглянуто існуючі рішення. Порівняно рішення задач, виділено недоліки існуючих рішень.

Проведено огляд існуючих аналогів систем, які є лідерами на ринку аналізу даних в банківській сфері, такі як SAS Enterprise, IBM SPSS, Deductor та Кредитний брокер Фокстрот. Сформульовано вимоги до системи, що розробляється. Наведено опис бізнес-процесів, сформульовано акторів та функції системи.

2 РОЗРОБКА РІШЕНЬ ДЛЯ ПОШУКУ КРЕДИТОРІВ В ТОРГІВЕЛЬНИХ МЕРЕЖАХ

2.1 Змістовна постановка задачі

Як було сказано в попередньому розділі, ринок потребує інформаційну систему-конфігуратор, яка б мала можливість взаємодії з банками. Для прискорення процесу пошуку і підбору кредитних продуктів, використовується прогнозування.

Для класифікації та прогнозування використовується скорингова модель на базі логістичної регресії. Для кожного банку будується своя скорингова модель.

Вхідними даними системи є:

- множина заявок попередніх клієнтів;
- множина відповідей банків на заявки клієнтів.

Заявка включає в себе множину характеристик клієнта.

На базі множини заявок і множини відповідей банків необхідно побудувати прогнозуючу скорингову модель, за допомогою якої можна прогнозувати відповіді банків в нових заявках.

2.2 Математична модель

Вхідними даними досліджуваної задачі є:

- A – множина заявок;
- $a_i \in A, i = \overline{1, n}$ – i -та заявка, де n – кількість заявок;
- B – множина відповідей банків на заявки;

- $b_i \in B, i = \overline{1, n}$ – відповідь банків i -ї заявки.
- $X_i = (x_{i,1}, x_{i,2} \dots x_{i,j}), i = \overline{1, n}, j = \overline{1, c}$, вектор значень параметрів заявки a_i , в якому $x_{i,j}$ значення j -го параметру заявки, де c – кількість характеристик клієнта;
- $L_i = (l_{i,1}, l_{i,2} \dots l_{i,m}), i = \overline{1, n}, m = \overline{1, v}$, вектор значень відповідей банків заявки b_i , в якому $l_{i,m}$ значення відповіді m -го банку, де v – кількість банків.

Вихідними даними досліджуваної задачі є:

- $k_{m,j}$ – ваговий коефіцієнт скорингової моделі m -го банку, j -ї ознаки;
- $S_{m,i}$ – скоринговий бал m -го банку, i -ї заявки.

Скоринговий бал визначається наступною формулою:

$$S_{m,i} = \sum_j k_{m,j} * x_{i,j}, \quad (2.1)$$

Основною задачею є пошук вагових коефіцієнтів $k_{m,j}$ для кожного m -го банку, j -ї ознаки. Тоді вихідне значення $S_{m,i}$ m -го банку, i -ї заявки будемо використовувати як прогноз відповіді від кожного банку на заявку клієнта.

Методи пошуку вагових коефіцієнтів і скорингових балів описані в підрозділі 2.4.

2.3 Огляд методів розв'язання

Для скорочення часу отримання відповіді на заявку клієнта та підвищення шансів клієнта отримати позитивну відповідь, потрібно використовувати скорингові моделі для прогнозування результатів відповідей банків. У разі невисоких шансів отримати кредит, формується рекомендація для клієнта змінити деякі параметри заявки для підвищення шансів отримання кредиту.

Щоб сформулювати відповідь клієнту, банки використовують скорингові моделі для оцінки платоспроможності клієнта позичальника. За допомогою скорингу можна отримати класифікаційну статистично-математичну модель по різних групах.

Для реалізації скорингової моделі використовують різні моделі. На ринку кредитування найчастіше використовуються такі методи: лінійний регресійний аналіз, логістичний регресивний аналіз, нейронні мережі та дерева рішень.

Для прогнозування дуже важливо мати реальний або генерований на базі реальних даних пул даних. При побудові банками скорингових моделей, більшість часу займає накопичення достатньої кількості даних на базі яких, банк зможе формувати ризики.

2.3.1 Лінійна ймовірнісна модель

Для аналізу даних двійкової відповіді зазвичай використовуються лінійні моделі в логістичному перетворенні ймовірностей. Інтерпретація даних у вигляді двійкових результатів виникає в багатьох галузях науки [5].

Лінійна ймовірнісна модель – це лінійна регресійна модель, залежні змінні якої набувають бінарного значення, тобто 0 або 1. Для описаної раніше предметної області бінарний результат підходить, він буде означати рекомендовано надати кредит чи ні. Формально дану модель можна описати так:

$$y = b_0 + b_1x_1 + b_2x_2 + \dots + b_nx_n, \quad (2.2)$$

де y – результуюча змінна, що в предметній області означає прогноз повернення кредиту позичальником;

b_0, \dots, b_m – коефіцієнти рівня лінійної регресії, що оцінюються для корегування значимості характеристик клієнта;

x_1, \dots, x_m – змінні що описують характеристики клієнта позичальника.

Однак у випадку лінійного аналізу ймовірностей існує кілька серйозних статистичних проблем.

По-перше, помилки бінарні. А це може призвести до неадекватності результатів, отриманих методом найменших квадратів.

По-друге, існує проблема з інтерпретацією результатів, оскільки знайдене значення ймовірності банкрутства може знаходитися поза діапазоном $[0, 1]$.

Хоча припущення, що лежать в основі лінійного аналізу ймовірності, відрізняються від припущень дискримінантного аналізу, результати дуже схожі. Цей факт частково пояснює відносну непопулярність вищезазначеного методу.

2.3.2 Логістична регресія

Логістична регресія – поширений метод для вирішення задач регресії і класифікації. В банківських скорингових системах цей метод є дуже поширеним для аналізу ризиків [5].

Результуюча змінна приймає лише два значення, що для предметної області означає був рекомендований кредит чи ні.

Ціллю логістичної регресії є аналіз зв'язків між цільовою функцією та змінними-параметрами. Так як у предметній області результат є бінарним, то і в даному випадку використовується бінарна версія моделі.

Логістична регресія надає змогу визначити ймовірність того, що буде отримано один з результатів.

Модель можна записати так:

$$y = F(x_1, \dots, x_n) \quad (2.3)$$

Функція F є лінійною функцією від змінних-параметрів, що можна описати рівнянням:

$$y = b_0 + b_1x_1 + b_2x_2 + \dots + b_nx_n, \quad (2.4)$$

Щоб вирішити цю проблему виходу функції за інтервал $[0,1]$, модель можна побудувати трохи по іншому. Ми відмовимось від бінарно цільової змінної, та будемо приймати значення в діапазоні $[0,1]$, незалежно від значень змінних-параметрів. Це реалізується завдяки використанню logit-перетворення:

$$p = \frac{1}{1 + e^{-y}}, \quad (2.5)$$

де p – ймовірність отримання певного результату;

y – значення регресійного рівняння.

По суті, логістична регресія надає не якийсь явний результат, а оцінює ймовірність отримання певного результату.

На рисунку 2.1 зображено залежність між величиною y та ймовірністю отримання певного результату.

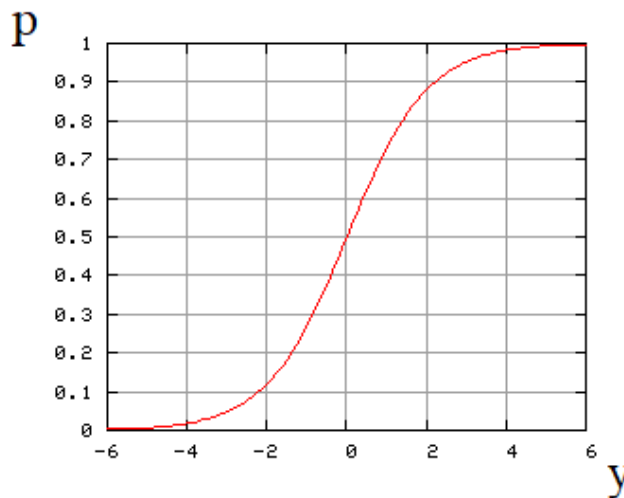


Рисунок 2.1 – Приклад логістичної кривої

2.4 Розробка алгоритму розв'язання

В результаті огляду методів вирішення задачі, вирішено використовувати метод логістичної регресії.

Найбільшою складністю в прогнозуванні методом логістичної регресії є пошук коефіцієнтів. Найбільш поширеним є метод максимальної правдоподібності.

2.4.1 Пошук коефіцієнтів логістичної регресії

Метод максимальної правдоподібності — це метод оцінювання невідомого параметра шляхом максимізації функції правдоподібності. Він ґрунтується на припущенні про те, що вся інформація про статистичну вибірку міститься у цій функції. Оцінка максимальної правдоподібності є популярним статистичним методом, який використовується для створення статистичної моделі на основі даних і забезпечення оцінки параметрів моделі. В інших джерелах його називають як метод найбільшої правдоподібності [20].

Для фіксованого набору даних і базової імовірнісної моделі, використовуючи метод максимальної правдоподібності, знайдемо значення параметрів моделі, які роблять дані «ближчими» до реальних. Оцінка максимальної правдоподібності дає унікальний і простий спосіб визначити рішення у разі нормального розподілу [20].

$$\hat{b} = \arg \max L(b), \quad (2.6)$$

де $L(b)$ – функція правдоподібності;

b – вектор параметрів;

\hat{b} – векторі оцінок параметрів.

Змінна y має розподіл Бернуллі, тому що може мати лише два значення 0 або 1. Знайдемо ймовірність того, що $y = 1$:

$$\Pr\{y = 1|x\} = f(z), \quad (2.7)$$

$$f(z) = \frac{1}{1 + e^{-z}}, \quad (2.8)$$

$$z = b_0 + b_1 x_1 + \dots + b_n x_n, \quad (2.9)$$

Тоді очевидно, що ймовірність того, що $y = 0$:

$$\Pr\{y = 0|x\} = 1 - f(z), \quad (2.10)$$

Тепер можна сформулювати функцію розподілу y при x . Вона буде мати такий вигляд:

$$\Pr\{y|x\} = f(b^T x)^y (1 - f(b^T x))^{1-y}, y \in \{0,1\}, \quad (2.11)$$

Після, можна на вибірці сформулювати функцію правдоподібності:

$$L(b) = \prod_{i=1}^n \Pr\{y = y^{(i)}|x = x^{(i)}\} \quad (2.12)$$

Прологарифмуємо функцію щоб потім її максимізувати:

$$\ln L(b) = \sum_{i=1}^n \left(y^{(i)} \ln f(b^T x^{(i)}) + (1 - y^{(i)}) \ln (1 - f(b^T x^{(i)})) \right), \quad (2.13)$$

Залишилось максимізувати дану функцію, зазвичай використовують метод градієнтного спуску. Ітераційно знайдемо максимум.

2.4.2 Скорингова карта

Скорингова карта – це множина ознак клієнта-позичальника (стать, вік, дохід, рухоме і нерухоме майно і т.д.). На базі вказаних в анкеті даних, формується скоринг бал. І порівнюючи цей скоринг бал, банк дає рішення.

Скоринг карта підходить як і бінарним ознакам (стать), так і значимим ознакам (дохід).

Основною задачею побудови скорингової таблиці є оцінка коефіцієнтів значущості ознак клієнтів. Підсумкову оцінку скорингу можна підрахувати за наступною формулою:

$$S = k_1x_1 + k_2x_2 + \dots + k_ix_i, \quad (2.14)$$

де k_i – ваговий коефіцієнт логістичної регресії i -ї ознаки клієнта;

x_i – значення i -ї ознаки.

2.4.3 Значущі ознаки

Зі всього масиву даних про клієнтів, потрібно обрати найбільш значущі ознаки. Зазвичай, оцінку ознаки визначають за допомогою WOE (Weight of Evidence). Уявімо, що є деякі події або явища, які, ймовірно, пов'язані з якимись іншими просторовими явищами. Нехай потрібно виявити закономірності - які саме характеристики впливають на ймовірність появи цієї події, наскільки сильний взаємозв'язок. Зрозуміло, що можна вирішити це завдання шляхом побудови рівняння регресії. Метод вагомості ознаки пропонує подібний підхід, але коефіцієнти, отримані за допомогою даного методу, легше інтерпретуються.

За своєю суттю метод вагомості ознаки представляє статистичний метод, заснований на теоремі Байеса. Розглядається найпростіший варіант методу -

бінарний, який має справу з двійковими категоріями даних, тобто коли шари містять інформацію, представлену двома класами (вологий / сухий, високий / низький, наявність / відсутність і т.п.).

Цей показник розраховується як:

$$W = \ln \left(\frac{d_i^{(1)}}{d_i^{(2)}} \right), \quad (2.15)$$

де $d_i^{(1)}$ - кількість або частка погоджених запитів на кредит по i -й ознаці клієнта;

$d_i^{(2)}$ - кількість або частка не погоджених запитів на кредит по i -й ознаці клієнта.

Тепер нехай деяка номінальна ознака під номером i приймає значення від 1 до n . Для цього потрібно розрахувати I (*Information Value*) за формулою:

$$I = \sum_{i=1}^k \left(d_i^{(1)} - d_i^{(2)} \right) * W_i, \quad (2.16)$$

де $d_i^{(1)}$ - частка погоджених запитів на кредит по i -й ознаці клієнта;

$d_i^{(2)}$ - частка не погоджених запитів на кредит по i -й ознаці клієнта;

W_i – значення Weight of Evidence для i -ї ознаки.

Для ознак з найбільшими значеннями I в одну групу об'єднуються категорії з близькими значеннями W . При такому підході важливо задати умови близькості значень W і кількість згрупованих категорій. Для подальшої зручності роботи з ознаками, їх можна поділити на групи по значимості. Групи ознак наведені в таблиці 2.1.

Таблиця 2.1 – Категорії значущості ознак

<i>Категорія значущості</i>	<i>Значення I</i>
Не значуща	$I < 0.01$
Низька	$0.01 < I < 0.2$
Середня	$0.2 < I < 0.4$
Висок	$0.4 < I < 0.6$
Дуже висока	$I > 0.6$

2.4.4 Категоріальні ознаки

Описані вище моделі розраховані на роботу з числовими змінними, але на практиці більшість даних про клієнта зберігається у вигляді категоріальних змінних. Наприклад ознака «Вид основного доходу». Варіантами цієї ознаки виступають наступні категорії:

- дохід СПД;
- заробітна платня (офіційне працевлаштування);
- заробітна платня (офіційне працевлаштування);
- пенсія;
- тимчасово не працює.

Для оцінки можна використовувати описане вище значення зваженої сукупності. W (Weight of evidence) показує значущу міру відповідної категорії ознаки позичальника.

$$W = \ln \left(\frac{d_i^{(1)}}{d_i^{(2)}} \right), \quad (2.17)$$

де $d_i^{(1)}$ - кількість або частка погоджених запитів на кредит по i -й ознаці клієнта;

$d_i^{(2)}$ - кількість або частка не погоджених запитів на кредит по i -й ознаці клієнта.

Використовувавши W кожній категорії можна проставити відповідне значення. Значення категорії наведено в таблиці 2.2.

Таблиця 2.2 Оцінка категоріальних змінних

<i>Категорія</i>	<i>W</i>
Дохід СПД	0.39
Заробітна платня (офіційне працевлаштування)	0.45
Заробітна платня (офіційне працевлаштування)	0.4
Пенсія	0.1
Тимчасово не працює	-0.3

2.5 Розробка методу розв'язання задачі

Реалізація системи складається з трьох ключових компонентів зображена на рисунку 2.2.

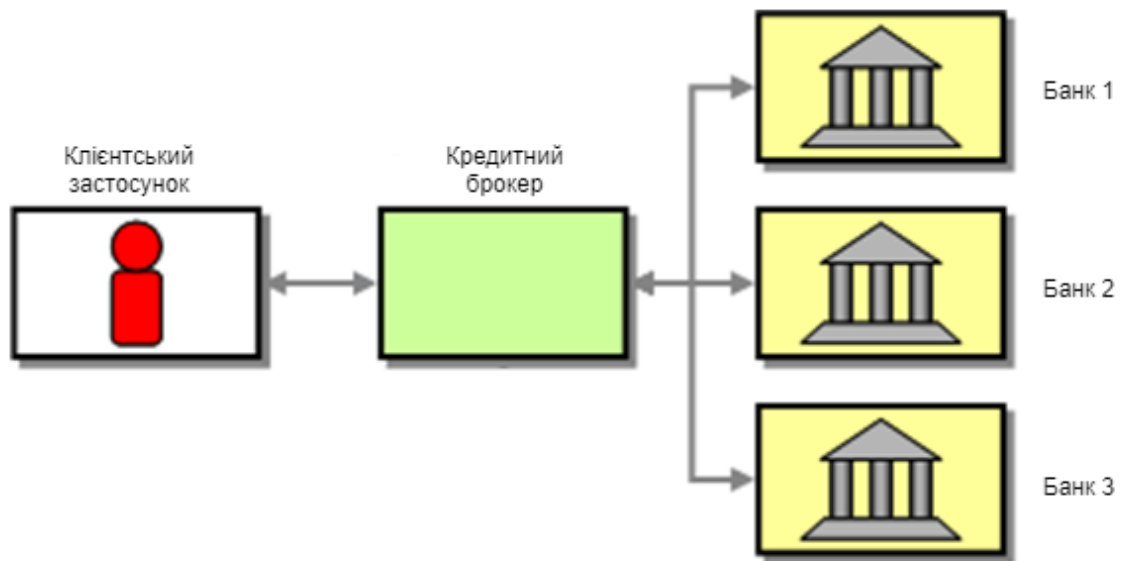


Рисунок 2.2 – Схема роботи кредитного брокера

- «Клієнтський застосунок» збирає і валідує дані;
- «Кредитний брокер» координує зв'язок між кредитним бюро і банками;
- кожен «Банк» отримує запити від «Кредитного брокера» та повертає результати роботи зовнішніх банківських скорингових систем.

У процесі роботи з банком, система не має доступу до внутрішніх скорингових систем банків. Система обмежена моніторингом та керуванням компонентами ззовні. Щоб розробити систему максимально універсальною, система ставитися до кожного з компонентів як до чорної скриньки. Маючи на увазі це обмеження, система має відповідати таким вимогам:

- має бути консоль управління - єдина панель, яка б відображала стан всіх компонентів;
- для тестування розроблено тестовий клієнт, який відстежує час відповідей брокера на позику між запитом та відповіддю, для оптимізації швидкості отримання відповідей від банків;

– періодично в банки будуть надсилатись тестові запити для перевірки працездатності банківських сервісів.

Консоль управління

Система повинна мати можливість збирати показники з декількох компонентів в одне місце, щоб можна було оцінити стан системи. Консоль управління також повинна мати можливість контролювати потік повідомлень та параметри компонентів, щоб можна було вирішити проблеми з перебоями шляхом перенаправлення повідомлень або зміни поведінки компонентів.

Консоль управління зв'язується з окремими компонентами через обмін повідомленнями. Вона використовує окрему шину управління, яка містить лише повідомлення, пов'язані з управлінням системою.

Дані на інтерфейсі будуть відображатись в режимі реального часу. Це можна реалізувати на базі таймерів, по яким буде опитуватись сервіс на зміну статусів. Є альтернативний варіант з використанням технології веб-сокетів, з використанням яких, інтерфейс в реальному часі буде реагувати на зміну статусу заявки.

Якість обслуговування кредитного брокера

Перша вимога до управлінського рішення - це вимірювання якості обслуговування, яке надає система своїм клієнтам. Для цього типу моніторингу інформативним є бізнес-вміст окремих повідомлень, тобто не анкетні дані про клієнта, а лише час, що минув між запитом та відповіддю. Складність відстеження часу між цими двома повідомленнями полягає в тому, що клієнт може вказати канал для відповіді через інший сервіс, у якого буде інша IP-адреса. Тому для відстеження таких запитів потрібно побудувати middle-wire систему, назовемо її Smart Proxy. Middle-wire система перехоплює повідомлення, зберігає повернені адреси для клієнта і замінює їх на адреси фіксованого каналу. В

результаті сервіс надсилає всі відповіді на один канал. Потім він пересилає відповідь на вихідну адресу повернення, вказану клієнтом. Схема зображена на рисунку 2.3.

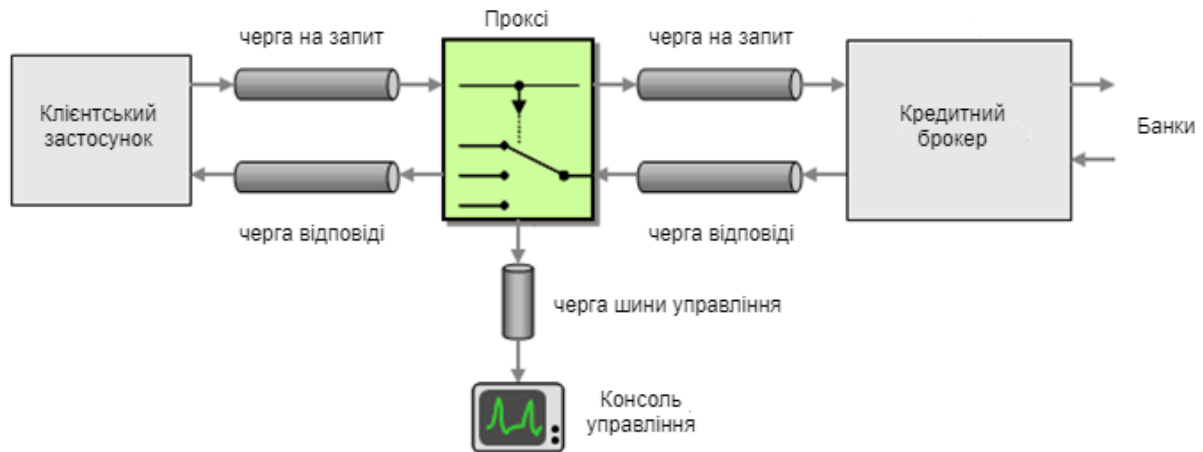


Рисунок 2.3 – Схема роботи Middle-ware системи

Для того, щоб скористатися Smart Proxy, його потрібно використовувати для підтримки зв'язку між клієнтом та брокером. Смарт проксі вказує брокеру, щоб він відправив всі повідомлення на `brokerReplyQueue` канал -чергу, звідки він передає повідомлення назад на правильну зворотну адресу, спочатку вказану клієнтом.

Smart Proxy використовується для вимірювання часу відповіді на запити на отримання позики та кількості запитів, які обробляються кредитним брокером в будь-який час. Smart Proxy може вимірювати час між запитом і відповідними повідомленнями та записати час коли повідомлення запиту було отримано для подальшої аналітики. При отриманні відповідного повідомлення відповіді, Smart Proxy віднімає час запиту від поточного часу для обчислення часу, що пройшов між запитом і відповіддю. Smart Proxy може оцінити, скільки активних запитів

брокер зберігає в черзі, тобто скільки в даний момент не відправлених запитів по різних причинах. Система може оновлювати кількість невідправлених і неотриманих запитів і повідомлень щоразу, коли отримуємо запит або відповідь.

Смарт проксі передає метрики на консоль управління для моніторингу та аналізу. Можна надсилати статистику для кожного окремого повідомлення, але це збільшує навантаження на систему, якщо іде справа про великі обсяги запитів і повідомлень. Вставлення Smart Proxy у потік повідомлень вже подвоює кількість відправлених повідомлень (два повідомлення запиту та відповіді замість одного на кожне), тому потрібно уникати надсилення іншого контрольного повідомлення для кожного повідомлення та запиту. Замість цього використовується таймер, щоб Smart Proxy надсилав метрики на шину управління через попередньо визначені інтервали, наприклад, кожні 5 секунд. Повідомлення про показники може містити або підсумкові показники (наприклад, максимальний, мінімальний та середній час відповіді), або детальну інформацію про всі повідомлення, які пройшли через інтервал. Для того, щоб дані з метриками були малими, а консоль управління простою, передаються просто зведені (прораховані) показники на консоль.

Перевірка роботи банку

Друга вимога до управління процесу кредитування - контролювати правильність роботи зовнішньої служби банків. Кредитний брокер отримує доступ до цієї послуги для отримання відповіді від банку, погодив він кредит чи ні.

Щоб перевірити правильність роботи зовнішньої служби банку, система надсилає періодичні тестові повідомлення до сервісу. Оскільки служба банку підтримує зворотну адресу, легко ввести тестове повідомлення без порушення потоку повідомлень. Схема системи моніторингу зображена на рисунку 2.4.

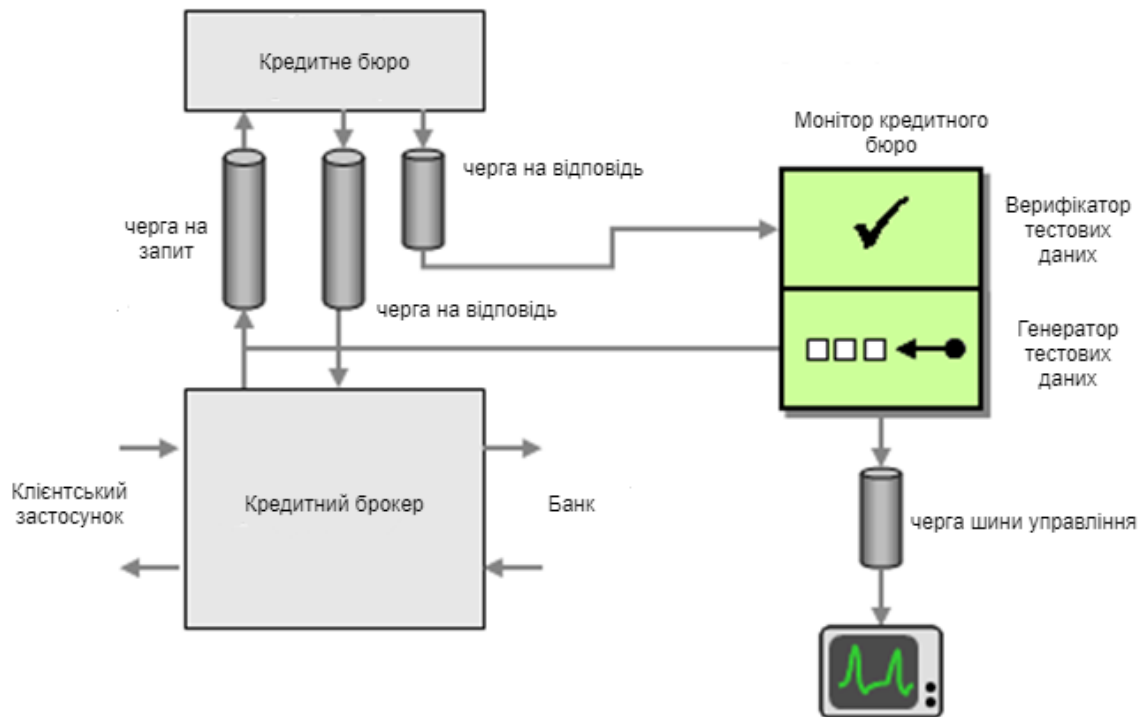


Рисунок 2.4 – Діаграма моніторингу сервісів банків

Для перевірки правильності роботи служби банку, потрібен генератор тестових даних та перевірка тест-даних. Генератор тестових даних створює тестові дані для надсилання до тестової служби. Для тестів використовується спеціальний, фіксований ІПН, який ідентифікує вигадану особу. Це дозволяє перевірити дані результату заздалегідь встановленими результатами. Таким чином можна не лише перевірити, чи було отримано відповідь, але й перевірити вміст повідомлення. В даному варіанті банк запрограмований для повернення випадкових результатів незалежно від ІПН, що надходить.

Підсистема перевірки тест-даних також перевіряє час відгуку зовнішньої служби. Якщо не було отримаємо відповідь у встановлений часовий проміжок, система сповіщає про це консолі управління. Щоб мінімізувати пропускну здатність мережі, сервіс повідомляє консоль лише у випадку, якщо відповідь затримана або неправильно сформована, а не коли служба працює належним

чином. Єдиний виняток із цього правила відбувається, коли монітор отримує правильну відповідь від служби після виявлення помилки. У такому випадку монітор надсилає на консоль управління повідомлення "ОК", щоб вказати, що банк знову працює правильно. Нарешті, під час запуску монітор надсилає повідомлення на консоль, щоб повідомити свій статус роботи. Це повідомлення дозволяє консолі виявити всі активні монітори, щоб вона могла відображати стан кожного з них.

Реалізація монітора використовує два окремих таймера. Таймер відправки визначає інтервал часу між останнім отриманим повідомленням або останньою подією очікування та відправленням наступного тестового повідомлення. Таймер тайм-аута запускається кожного разу, коли монітор надсилає повідомлення із запитом. Якщо відповідь надходить у вказаний інтервал очікування, таймер тайм-ауту скидається та перезавантажується наступним запитом. Якщо монітор не отримує відповідь на повідомлення протягом зазначеного інтервалу, таймер тайм-ауту зупиняється і монітор надсилає повідомлення про помилку на шину управління. Потім запускається новий таймер відправки, щоб ініціювати нове повідомлення запиту після інтервалу надсилання. Діаграма моніторингу служб банків зображена на рисунку 2.5.



Рисунок 2.5 – Діаграма моніторингу служб банків

Відмова одного з банківських сервісів

Тепер, коли система може відстежувати стан зовнішньої служби банку, потрібно використовувати ці дані для реалізації схеми відмови, щоб брокер міг продовжувати працювати навіть тоді, коли служба банку не працює. Коли використовується кілька конкуруючих споживачів на одному каналі «точка-точка», відмова одного споживача не перерве обробку до тих пір, поки інші споживачі все ще працюють. Коли багато споживачів активні, вони розподіляють навантаження, ефективно застосовуючи простий механізм балансування навантаження.

В розділі 2.5 наведено підмножину основних функцій управління системою, які можна реалізувати за даною схемою.

2.6 Результати досліджень ефективності побудованої скорингової моделі в прогнозуванні

2.6.1 Оцінка скорингових моделей

Після того як скорингова модель побудована, необхідно визначити якість створеної скорингової системи, тобто провести верифікацію скорингового балу. Один з найбільш поширених способів верифікації скорингових систем - аналіз ROC-кривої (англ. Receiver Operating Characteristic).

Загально прийнято використовувати для оцінки якості класифікації скорингових моделей площа під ROC-кривою (AUC), яка представляє собою міру здатності моделі розрізняти «хороших» і «поганих» позичальників. У разі якщо необхідно порівняти між собою скорингові моделі, зазвичай порівнюють їх AUC: модель, у якій цей показник найбільший, має кращу розділяє здатність і, отже, є кращою. Однак не завжди застосування скорингової моделі з найбільшою AUC доцільно для банку з точки зору прибутковості. Метою нашого

дослідження є розробка показника якості дискримінації скорингової моделі, що враховує не тільки здатність моделі відокремлювати дисциплінованих позичальників від неплатників, але і більш вигідних позичальників від менш вигідних.

Банки застосовують скорингові моделі в процесі роздрібного кредитування з метою виявлення потенційно більш ризикових позичальників і скорочення втрат, викликаних відмовою боржників платити за своїми зобов'язаннями. Чим точніше скорингова модель розділяє «хороших» і «поганих» позичальників, тим менше втрати і більший прибуток банку.

Якщо позичальник є «поганим» і класифікується моделлю як «поганий», він вважається істинно позитивним (TP); якщо він класифікується як «хороший», то вважається помилково негативним (FN). Якщо позичальник є «хорошим» і класифікується як «хороший», він вважається істинно негативним (TN); якщо він класифікується як «поганий», це вважається помилковим спрацьовуванням (FP). Точність моделі характеризується двома показниками: чутливістю (Sensitivity) і специфічністю (Specificity). Чутливість вимірюється часткою істинно позитивних результатів - часткою позитивних результатів у «поганих» позичальників, розраховується за формулою $TP / (TP + FN)$. Специфічність вимірюється часткою істинно негативних результатів у «хороших» позичальників, розраховується за формулою $TN / (TN + FP)$. Пари значень чутливості і специфічності на всьому діапазоні відповідають координатам ROC-кривої.

Нехай є вибірка позичальників $S = \{(p_1, y_1), (p_2, y_2), \dots, (p_n, y_n)\}$, де $p_i \in \mathbb{R}$ - ймовірність приналежності позичальника до класу «поганих», n – кількість позичальників, розрахована на основі скорингової моделі; $y_i \in \{0, 1\}$ - фактична приналежність позичальника до класу «поганих», якщо $y = 1$, або «хороших», якщо $y = 0$. Тоді площа під ROC-кривою, AUC, для даної моделі може бути обчислена за формулою:

$$AUC = \frac{\sum s(p_i, p_j)}{n^0 * n^1}, \quad (2.18)$$

де n^0 і n^1 кількість позичальників в вибірці, що відносяться до класу «хороших» і «поганих» відповідно, і $s(p_i, p_j)$ задається таким чином:

$$s(p_i, p_j) = \begin{cases} 0, & \text{якщо } p_i > p_j \\ 0.5, & \text{якщо } p_i = p_j \\ 1, & \text{якщо } p_i < p_j \end{cases} \quad (2.19)$$

Для трансформації значення в AUC в значення в діапазоні від 0 до 1 потрібно перетворити його в коефіцієнт G за формулою:

$$G = 2 * AUC - 1. \quad (2.20)$$

В даному випадку при розрахунку AUC позичальники поділяються на два класи тільки за однією ознакою - ймовірності дефолту. Якщо видавати всім позичальникам кредит за однією ставкою, то порівняння моделей на основі розрахованої таким чином AUC буде справедливим. Однак на практиці в банках діє система індивідуальних знижок і надбавок, тому пропоновані позичальникам ставки розрізняються. Звідси виникає проблема вибору скорингової моделі, застосування якої дозволить отримати найбільший прибуток.

Нехай порівнюються дві скорингові моделі, які на основі соціальних характеристик відносять позичальника до класу «поганих». Кожному позичальнику ставиться у відповідність три параметра: p_j - ймовірність дефолту, розрахована на основі скорингової моделі, r_j - ефективна процентна ставка і y_j - індикатор фактичної приналежності позичальника до класу «поганих». Трансферна ставка для банку дорівнює r_{tr} .

Прибуток банку на період в 1 рік розраховується як сума фінансових результатів за всіма договорами. Фінансовий результат за j -м договором, отриманий за підсумками року, розраховується за формулою.

$$C_j = \begin{cases} D_j * (r_j - r_{tr}), & \text{якщо } y_j = 0 \\ -D_j, & \text{якщо } y_j = 1 \end{cases} \quad (2.21)$$

де D_j – сума видачі за j -м договором.

Отриманий показник має сенс використовувати для порівняння скорингових моделей, якщо в банку не реалізовано ціноутворення на основі індивідуальної оцінки кредитного ризику позичальника.

2.6.2 Аналіз результатів

Розглянемо приклад прогнозування кредитоспроможності позичальника по розробленій скоринг моделі. Для аналізу згенеруємо чотири анкети, які б описували різних клієнтів. В таблиці 2.3 наведено основні характеристики клієнтів.

Таблиця 2.3 – Анкетні дані клієнтів

<i>Ознака</i>	<i>Анкета 1</i>	<i>Анкета 2</i>	<i>Анкета 3</i>	<i>Анкета 4</i>
<i>Прострочені платежі по кредитах за рік</i>	Ні	Ні	Так	Ні
<i>Відсутність пропуску платежів</i>	Так	Так	Ні	Так
<i>Дохід</i>	32000	7000	16000	8000
<i>Стаж роботи</i>	6 років	1 рік	17 років	Не працює
<i>Вік</i>	27	19	45	67
<i>Освіта</i>	Вища	Незакінчена вища	Вища	Середня спеціальна
<i>Є діти?</i>	Ні	Ні	Так	Ні

Для прогнозування кредитоспроможності клієнтів потрібно оцінити вагові коефіцієнти кожної i -ої ознаки для кожного m -го банку. По описаному в попередньому підрозділі методу оцінки вагових коефіцієнтів, проведемо обчислення. В таблиці 2.4 наведено вагові коефіцієнти для трьох умовних банків.

Таблиця 2.4 – Вагові коефіцієнти 3-ьох банків

<i>Ознака</i>	<i>Вага ознаки банку 1</i>	<i>Вага ознаки банку 2</i>	<i>Вага ознаки банку 3</i>
<i>Прострочені платежі</i>	-0,8	-0,5	-0,8
<i>Відсутність пропуску платежів за рік</i>	0,7	0,9	0,5
<i>Дохід</i>	1	1,2	0,8
<i>Стаж роботи</i>	0,5	0,5	0,6
<i>Вік</i>	0,7	0,6	0,7
<i>Освіта</i>	0,7	0,6	0,6
<i>Є діти?</i>	-0,5	-0,2	-0,4

Перед використанням даних в анкеті, їх потрібно категоризувати. Методи категоризації описані в попередньому підрозділі. В таблиці 2.5 наведено категоризовані дані клієнтів.

Таблиця 2.5 – Категоризовані дані

Ознака	Анкета 1	Анкета 2	Анкета 3	Анкета 4
<i>Прострочені платежі</i>	0	0	1	0
<i>Відсутність пропуску платежів за рік</i>	1	1	0	1
<i>Дохід</i>	1	0,3	0,7	0,3
<i>Стаж роботи</i>	0,5	0,3	1	0
<i>Вік</i>	0,7	0,2	1	0,1
<i>Освіта</i>	1	0,3	1	0,5
<i>Є діти?</i>	0	0	1	0

Дані готові для побудови скорингової моделі. Візьмемо чотири анкети і визначимо скоринговий бал для першого банку (таблиця 2.6).

Таблиця 2.6 – Скорингова модель для кожної анкети

Ознака	Вага ознаки	Анкета 1	Анкета 2	Анкета 3	Анкета 4
<i>Прострочені платежі</i>	-0,8	0	0	1	0
<i>Відсутність пропуску платежів за рік</i>	0,7	1	1	0	1
<i>Дохід</i>	1	1	0,3	0,7	0,3
<i>Стаж роботи</i>	0,5	0,5	0,3	1	0
<i>Вік</i>	0,7	0,7	0,2	1	0,1
<i>Освіта</i>	0,7	1	0,3	1	0,5
<i>Є діти?</i>	-0,5	0	0	1	0
<i>Скоринг бал</i>		3,14	1,5	1,3	1,42
<i>Відповідь банк</i>		Погоджено	Погоджено	Відмовлено	Відмовлено

Для обраного банку осередненим скоринг балом, при якому банк дає позитивну відповідь є **2,05**.

По анкеті першого клієнта видно, що він є прикладом «хорошого» клієнта для кредитора. Скоринг бал це підтверджує, тому що значно перевище осереднений бал, тому скорингова модель спрогнозувала відповідь вірно.

По другій анкеті видно, що це «сірий» клієнт. Скоринг бал це підтверджує. Отриманий скоринг бал менше прохідного, проте цей клієнт вже був клієнтом обраного банку, тому банк дав позитивну відповідь.

Третя і четверта анкета є прикладом «поганих» клієнтів. Скоринг бал це підтвердив і вірно спрогнозував відмову банку.

На прикладі першої анкети сформуємо модель в розрізі кожного банку. Модель представлена в таблиці 2.7.

Таблиця 2.7 – Скорингова модель для анкети в розрізі банків

<i>Ознака</i>	<i>Параметри анкети 1</i>	<i>Вага ознаки банку 1</i>	<i>Вага ознаки банку 2</i>	<i>Вага ознаки банку 3</i>
<i>Прострочені платежі</i>	0	-0,8	-0,5	-0,8
<i>Відсутність пропуску платежів за рік</i>	1	0,7	0,9	0,5
<i>Дохід</i>	1	1	1,2	0,8
<i>Стаж роботи</i>	0,5	0,5	0,5	0,6
<i>Вік</i>	0,7	0,7	0,6	0,7
<i>Освіта</i>	1	0,7	0,6	0,6
<i>Є діти?</i>	0	-0,5	-0,2	-0,4
<i>Скоринг бал</i>		3,14	3,37	2,69
<i>Усереднений бал надання позитивної відповіді</i>		2,05	2,8	3,2
<i>Відповідь банку</i>		Так	Так	Ні

Скорингова модель, очевидно, спрогнозувала позитивну відповідь у першого і другого банку.

Оцінимо якість побудованих скорингових моделей для кожного банку. Для оцінки знайдемо загальну точність та коефіцієнт G . Оцінка буде проводитись на 1000 випадкових заявках. В таблиці 2.8 наведено отримані результати.

Таблиця 2.8 – Коефіцієнти точності моделі

<i>Ознака</i>	<i>Банку 1</i>	<i>Банку 2</i>	<i>Банку 3</i>
<i>Загальна точність</i>	<i>0,695</i>	<i>0,728</i>	<i>0,715</i>
<i>Коефіцієнт G</i>	<i>0,759467</i>	<i>0,782187</i>	<i>0,763542</i>

Висновок до розділу

В розділі розглянуто існуючі методи прогнозування, за допомогою яких можна оцінювати платоспроможність клієнтів-позичальників.

Розглянуто методи лінійно ймовірнісної моделі, такі як логістична регресія та скорингова карта. Побудовано скорингові моделі для оцінки шансів отримання позитивної відповіді від банку. Наведено приклади застосування моделей.

Розглянуто способи оцінки скорингових моделей. Проаналізовано отримані результати та показано, що розроблена схема вирішення задачі може застосовуватись в системах для підбору кредитних пропозицій.

В розділі також розглянуто реалізацію інформаційно-пошукової системи. Представлено схеми, по яким застосунок буде передавати дані банку, та отримувати від них відповіді. Описано загальну схему роботи програмного забезпечення.

3 ОПИС ПРОГРАМНОГО ТА ТЕХНІЧНОГО ЗАБЕЗПЕЧЕННЯ

3.1 Засоби розробки

При створенні програмного продукту були використанні оболонки JetBrains Rider для програмування на мові C#, JetBrains WebStorm для програмування на мові TypeScript та CSS та JetBrains DataGrip для керування базою даних SQL Server.

Для розробки Back-end частини було обрано мову програмування C# через приналежність до платформи .NET та мультиплатформеної платформи .NET Core. За допомогою .NET Core є можливість вести розробку застосунків на різних операційних системах, таких як: Windows, Linux та macOS. Цією перевагою вигідно виділяється .NET Core перед .NET, тому що з .NET можна працювати лише в операційній системі Windows. Платформи .NET та .NET Core включають в себе технологію розробки web-застосунків ASP.NET. У випадку ASP.NET, веб-API – це простий спосіб реалізації веб-служб RESTful застосунків за допомогою .NET framework. Веб-сервіси RESTful – це ті, що використовують HTTP як основний метод комунікації. Web-API ASP.NET зазвичай використовується як фреймворк, який дає можливість розробляти HTTP-служби, які б охопили об'єкти клієнта, такі як браузер, пристрої або планшети. ASP.NET Web API можна використовувати з MVC для будь-якого типу додатків. Таким чином, .NET веб-API є дуже важливим для розробки веб-додатків ASP.NET.

Платформа .NET Core підтримує бібліотеку SignalR, яка використовується для трансляції даних на інтерфейс в реальному часі. ASP.NET SignalR – це бібліотека для розробників ASP.NET, що спрощує процес додавання веб-функціональних можливостей у реальному часі до додатків. Веб-функціональність у режимі реального часу – це можливість мати серверний код, який набирає контент для підключених клієнтів, як тільки він стане доступним,

замість того, щоб сервер чекав, коли клієнт запитає нові дані. SignalR може використовуватися для додавання будь-якого типу «реального часу» веб-функціональності до вашої програми ASP.NET. Кожного разу, коли користувач оновлює веб-сторінку, щоб побачити нові дані, або сторінка впроваджує довгі опитування для отримання нових даних, вона є кандидатом для використання SignalR. Приклади включають приладові панелі та програми моніторингу, спільні програми (наприклад, одночасне редагування документів), оновлення прогресу роботи та форми в реальному часі.

C# використовує об'єктно-орієнтований підхід до програмування. Це означає, що потрібно описувати абстрактні конструкції на основі предметної області, а потім реалізувати між ними взаємодію. Також у мові активно використовується так званий «синтаксична цукор», який робить дуже спрощує програмування і збільшує швидкість написання коду. Замість того, щоб вручну писати багато коду, можна використовувати готову конструкцію. Деякі такі конструкції не є найкращими з точки зору продуктивності. Але все це перекривається за високої швидкості розробки.

Було обрано базу даних SQL Server для збереження даних. SQL Server - система управління реляційними базами даних (РСУБД), розроблена корпорацією Microsoft. Мовою запитів є Transact-SQL, створена Microsoft і Sybase. Transact-SQL є реалізацією стандартів ANSI / ISO за структурованою мовою запитів (SQL). Використовується для роботи з базами різних розмірностей. Конкурує з іншими СУБД у цьому сегменті ринку.

Клієнтська частина web-застосунку розроблена за допомогою ASP.NET. Фреймворк дозволяє з легкістю зв'язати Back-end і Front-end частини. Для покращення зовнішнього вигляду застосунку, було вирішено використовувати бібліотеку CSS ситтлів Bootstrap. Вона дозволяє налаштувати grid сторінки та включає в себе багато готових рішень для дизайну сторінок.

Back-end частина розроблялася на платформі JetBrains Rider. Rider включає в себе зручний редактор коду, що підтримує IntelliSense, а також авто-рефакторинг коду. IntelliSense – це утиліта, яка підказує розробнику можливі варіанти завершення кодових команд. В платформу інтегрований відладчик, що дуже спрощує пошук помилок. Платформа включає в себе можливість трасувати виконання коду. Цей інструментарій дозволяє знаходити слабкі місця в програмній реалізації та оптимізувати виконання коду. Rider повністю підтримує GIT і надає зручний інтерфейс для користування системою контролю версій.

Front-end частина розроблялася на кроссплатформенному редакторі коду JetBrains WebStorm. У своєму сервісі WebStorm має редактор вихідних кодів. WebStorm збільшує продуктивність програмування, завдяки підтримці сотень мов, підсвічуванню синтаксису, перевірки відповідності дужок, автоматичного відступу і багато іншого. Зрозумілі і налаштовувані комбінації клавіш, легкі налаштування та доповнення клавіатурних скорочень дозволяють легко переміщатися по коду.

3.2 Вимоги до технічного забезпечення

3.2.1 Загальні вимоги

Програмне забезпечення складається з серверної частини і клієнтської частини. Серверна і клієнтська частина розгортається на сервері.

Для правильної роботи даної програми до складу технічних засобів повинні входити:

- комп'ютер з такою конфігурацією:

- а) процесор з частотою не нижче 2,2 ГГц;
- б) об'єм RAM (не менше 8 ГБ);
- в) інші складові не так сильно впливають на роботу системи.

- комп'ютерна периферія, до складу якої входить:

- а) мишка;
- б) монітор;
- в) клавіатура;
- г) Wi-Fi роутер.

Висновок до розділу

В даному розділі було розглянуто засоби розробки та платформи, на яких було розроблене програмне забезпечення. Було задано загальні вимоги до технічного забезпечення.

4 РОЗРОБЛЕННЯ СТАРТАП-ПРОЕКТУ

На сьогодні на ринку банківських послуг дуже велика конкуренція, а вимоги клієнтів тільки ростуть. Споживачі, перед тим, як обирати кредитний продукт, хочуть порівняти між собою усі пропозиції на ринку та обрати найкращий серед запропонованих кредитних продуктів. Це зробити досить складно, потрібно відвідати кожен банк або переглянути відповідний веб-сайт банку, щоб зрозуміти, що він може запропонувати.

Особливо гостро ця проблема стоїть в сфері продажу товарів та послуг під споживчий кредит або так звану розстрочку. В Україні та у світі дуже мало торгових мереж, які можуть особисто видавати такі споживчі кредити. Загалом, магазини, в таких випадках, користуються послугами банків, проте без налагодженої інфраструктури, співпрацювати з більше ніж одним банком дуже складно і через це, з часом, обраний банк стає монополістом і починає завищувати процентні ставки на споживчий кредит. В такому випадку магазини починають втрачати потенційних клієнтів і гроші.

Також є проблема швидкості видачі продукті. Клієнт хоче прийти в магазин і максимум за 30 хвилин отримати кредитний продукт на купівлю обраного товару без відвідування банку.

У даному дослідженні розглядаються модель продажу кредитних продуктів у автоматичному режимі, без безпосередньої участі банку, використовуючи інформаційні технології. Пропонується створити інформаційну систему, яка дозволяє переглядати запропоновані кредитні продукти від усіх підключених до системи банків в одному місці. Система не тільки підбиратиме кредитні продукти для клієнта, а і в онлайн режимі надсилатиме запити на отримання кредиту в усі банки по вибраним кредитним продуктам.

4.1 Опис ідеї проекту

У таблиці 4.1 представлено зміст ідеї стартап-проекту, можливі області застосування і основні переваги, які може отримати користувач продукту.

Таблиця 4.1 – Опис ідеї стартап проекту

Зміст ідеї	Напрямок застосування	Вигода для користувача
Програмне забезпечення для підбору кредитних пропозицій клієнтам торговельних мереж	Торгові мережі, банки	Дозволяє торговим мережам продавати більше за рахунок збільшення кількості погоджених кредитів, та дозволяє банкам продавати більше кредитних продуктів

4.2 Аналіз ринку

Виявлення ринкових можливостей, які можуть бути використані при реалізації проекту на ринку, і ринкових загроз, які можуть перешкодити реалізації проекту, дозволяє планувати напрямки розвитку проекту з урахуванням ринкового середовища, потреб потенційних клієнтів і пропозицій конкуруючих проектів. Характеристика ринку стартапу описані в таблиці 4.2.

Таблиця 4.2 – Характеристика ринку стартапу

№ п/п	Показники стану ринку	Характеристика
1	Кількість конкурентів	2
2	Обсяг продаж в квартал	45 000 ум.од
3	Динаміка ринку	Зростає

№ п/п	Показники стану ринку	Характеристика
4	Наявність обмежень до входу	Немає
5	Специфічні вимоги до стандартизації та сертифікації	Немає
6	Середня норма рентабельності в галузі	55%

Проаналізувавши ринок, можна сказати, що ринок є сприятливим для продукту.

Визначимо потенційні групи клієнтів, їх характеристики та сформуємо орієнтований список вимог до продукції для кожної групи. Характеристики потенційних клієнтів стартап-проекту описані в таблиці 4.3.

Таблиця 4.3 – Характеристика потенційних клієнтів стартап-проекту

№ п/п	Потреба	Цільова аудиторія	Відмінності у поведінці груп клієнтів	Вимоги споживачів
1	Пошук кредитних пропозицій для клієнтів торгових мереж	Торгові мережі і банки	Застосунок використовує одна група клієнтів	Висока швидкість пошуку кредитних пропозицій

Проаналізуємо ринкову середу: таблиці факторів, здатних до ринкової реалізації проекту та факторів, які цим застосовують. Фактори загроз описані в таблиці 4.4. Фактори можливостей описані в таблиці 4.4.

Таблиця 4.4 – Фактори загроз

№ п/п	Фактор загрози	Зміст загрози	Дії компанії
1	Велика конкуренція	Прихід на ринок компанії з високим фінансуванням	Спланувати вихід з ринку. Старі партнери деякий час будуть використовувати програмне забезпечення, але з часом перейдуть до конкурента.
2	Політика	Зміна законодавства або зміна ідеології в країні	Перехід на ринок сусідніх країн.

Таблиця 4.5 – Фактори можливостей

№ п/п	Фактор можливості	Зміст	Дії компанії
1	Витіснення конкурентів	Витіснення конкурентів з менш якісним програмним забезпеченням	Розширення мережі партнерів. Заохочення партнерів конкурентів.
2	Розширення груп кредитування	Інтегрування кредитування в нові ринки	Вкладення ресурсів на пошук нових ринків продажу кредитних продуктів

Проаналізуємо пропозиції: знайдемо риси конкуренції на ринку.
 Ступеневий аналіз конкуренції на ринку описаний в таблиці 4.6.

Таблиця 4.6 – Ступеневий аналіз конкуренції на ринку

Особливості конкурентного середовища	В чому проявляється дана характеристика	Вплив на діяльність
1. Вказати тип конкуренції - монополістична	Існує один гравець монополіст і кілька конкурентів	Покращення продукту
2. За рівнем конкурентної боротьби - локальний	Фірми конкуренти з однієї країни	Підтримка продукту на внутрішньому ринку
3. За галузевою ознакою - внутрішньогалузева	Програмне забезпечення використовується в одній галузі	Модифікація застосунку для використання в інших галузях
4. Конкуренція за видами товарів - товарно-родова	Наявна конкуренція	Покращувати та розшифрувати функціонал та акцентувати на перевагах застосунку
5. За характером конкурентних переваг - нецінова	Покращення якості програмного застосунку	Розробляти нові застосунки схожого типу, але для іншого ринку
6. За інтенсивністю – не марочна	Торгова марка відома лише всередині галузі	Рекламувати бренд компанії всередині галузі

Визначимо та сформуємо перелік факторів конкурентоспроможності. Фактори конкурентоспроможності описані в таблиці 4.7.

Таблиця 4.7 – Фактори конкурентоспроможності

№ п/п	Фактор конкурентоспроможності	Обґрунтування
1	Ціна	Застосунок номінально безкоштовний, береться лише комісія за виданий кредит
2	Функціонал	Широкий функціонал програмного забезпечення, висока масштабованість дозволяє легко додавати новий функціонал
3	Дизайн та інтерфейс	Сучасний та швидкий інтерфейс якісно відрізняє застосунок від конкурентів

Проведемо SWOT-аналіз. SWOT — аналіз дозволяє провести детальне дослідження зовнішнього й внутрішнього середовища. Результатом раціонального SWOT-аналізу, спрямованого на формування загальненого інформаційного потенціалу, повинні з'явитися ефективні рішення, що стосуються відповідної реакції (впливу) суб'єкта (слабкої, середньої й сильної) відповідно до сигналу (слабкому, середньому або сильному) зовнішнього середовища.

Відмінна риса розглянутого підходу до проведення SWOT-аналізу на підприємстві полягає у такому:

- його побудова базується на методології системно-цільового підходу, де основна увага акцентується на вимірюванні параметрів зовнішнього й внутрішнього середовища в просторі, у часі й з урахуванням інформаційного потенціалу;
- проведення структуризації факторів зовнішнього й внутрішнього середовища, що є універсальними для будь-якого підприємства;
- здійснення синтезу факторів зовнішнього й внутрішнього середовища, що далі, в перспективі, відобразиться у системі рішень.

Сильні сторони:

- ціна;
- мікро-сервісна архітектура;
- сучасний та швидкий інтерфейс.

Слабкі сторони:

- складність розгортання застосунку.

Можливості:

- конкурентно спроможний ринок;
- постійна поява нових користувачів.

Загрози:

- поява конкурента гіганта.

На основі SWOT-аналізу визначимо варіанти ринкової поведінки для виведення стартап-проекту на ринок і оптимальний час їх реалізації на ринку з урахуванням потенційних конкурентів, які можуть бути виведені на ринок. Альтернативи ринкового впровадження описані в таблиці 4.8.

Таблиця 4.8 – Альтернативи ринкового впровадження

№ п/п	Альтернатива поведінки	Ймовірність отримання результатів	Строки реалізації
1	Створення Web API	90%	6 міс
2	Створення Front- end частини	95%	8 міс
3	Інтеграція банків і партнерів	60%	6 міс

4.3 Розроблення ринкової стратегії

Спочатку потрібно визначити стратегії охоплення ринку, описати цільові групи потенційних клієнтів описані в таблиці 4.9.

Таблиця 4.9 – Вибір цільових груп потенційних споживачів

№ п/п	Опис профілю цільової групи	Готовність споживачів сприйняти продукт	Орієнтований попит в межах цільової групи	Інтенсивність конкуренції в сегменті	Простота входу в сегмент
1	Торгові мережі	Висока	Високий	Висока	Низька складність
2	Банки	Середня	Середній	Низька	Середня складність
Які цільові групи обрано: 1					

Потрібно сформувати базову стратегію розвитку для обраних сегментів ринку описані в таблиці 4.10.

Таблиця 4.10 – Базові стратегії розвитку

№ п/п	Альтернатива розвитку	Стратегія охоплення ринку	Ключові конкурентоспроможні позиції	Базова стратегія розвитку
1	Акцентування уваги на унікальних функціях продукту, які виділяють товар на фоні конкурентів	Задати потреби кожної цільової групи та побудувати стратегію заохочування користувачів	Швидке реагування на зміни в ринку, якість продукту	Стратегія збільшення долі ринку

Задамо стратегію конкурентної поведінки. Базові стратегії конкурентної поведінки описані в таблиці 4.11.

Таблиця 4.11 – Базові стратегії конкурентної поведінки

№ п/п	Чи є проект «першопрохідцем » на ринку?	Чи буде компанія шукати нових споживачів, або забирати існуючих у конкурентів?	Чи буде компанія копіювати основні характеристики товару конкурента, і які?	Стратегія конкурентної поведінки
1	Ні	Змішана стратегія	Ні	Стратегія збільшення долі ринку

Задамо ринкову позицію, за якою споживачі будуть розрізняти проект серед конкурентів. Стратегії позиціонування описані в таблиці 4.12.

Таблиця 4.12 – Стратегії позиціонування

№ п/п	Вимоги	Стратегія розвитку	Конкуренто- спроможні позиції проекту	Вибір асоціацій, які мають сформувану комплексну позицію власного проекту
1	Швидкий та інтуїтивний клієнтський застосунок, безпечність використання	Стратегія заняття конкурентної ніші	Виділення ключових відмінностей продукту	Збільшення прибутку, зменшення ризиків

4.4 Побудова маркетингової програми

Підведемо підсумки попереднього аналізу конкурентоспроможності продукту. Основі переваги товару описані в таблиці 4.13.

Таблиця 4.13 – Основі переваги товару

№ п/п	Потреба	Вигода, яку надає товар	Перевага перед конкурентом
1	Збільшення продажів	Продукт надає можливість потенційним клієнтам торгових мереж купити товар, навіть при відсутності коштів	Автоматизована система зв'язку з банком
2	Сучасний інтерфейс	Сучасний інтерфейс збільшує швидкість заведення заявки	Багаторівнева анкета
3	Надійність системи	Система буде працювати, навіть якщо один з банків не буде на зв'язку	Система безперебійної роботи

Задамо маркетингову модель. Рівні моделі товару описані в таблиці 4.14.

Таблиця 4.14 – Рівні моделі товару

Рівні	Сутність та складові продукту		
Товар за задумом	Програмне забезпечення для пошуку та підбору кредитних продуктів в торгівельних мережах		
Товар у реальній реалізації	Властивості	М/Нм	Вр/Тх/Тл/Е/Ор
	1. Збір і мапінг даних; 2. Побудова прогнозуючої скорингової моделі; 3. Запит до банку		
	Якість: автоматичне тестування усіх важливих етапів		
	Пакування: відсутнє		
	Марка: «Untitled»		
Товар із підкріпленням	До продажу: відсутнє		
	Після продажу: навчання, суппорт		
Вихідний код не є відкритим			

Потрібно встановити цінові межі, щоб становити ціну на товар. Межі встановлення цін описані в таблиці 4.15.

Таблиця 4.15 – Межі встановлення цін

№ п/п	Ціна на продукти замітники	Ціна на продукти аналоги	Рівень доходів споживачів	Верхня та нижня межі встановлення ціни на товар/послугу
1	\$1500	\$1000	Середній рівень доходів	Покупка та інтеграція: нижня межа - \$800, верхня межа - \$1500

Задамо оптимальну систему збуту. Система збуту описана в таблиці 4.16.

Таблиця 4.16 – Система збуту

№ п/п	Специфіка закупівельної поведінки цільових клієнтів	Функції збуту, які має виконувати постачальник товару	Глибина каналу збуту	Оптимальна система збуту
1	Клієнти – торгові мережі, які бажають інтегрувати в свій процес продажу товарів, продажу товарів під кредит в автоматичному режимі. Вони готові піти на інтеграцію заради збільшення кількості продажів.	Пропонування послуг потенційним клієнтам.	Однорівнева глибина збуту, послуги надаються напрямку від компанії	Прямий канал збуту

Останньою складовою маркетингової програми є розроблення концепції маркетингових комунікацій, що спирається на попередньо обрану основу для позиціонування, визначену специфіку поведінки клієнтів. Концепція маркетингових комунікацій описана в таблиці 4.17.

Таблиця 4.17 – Концепція маркетингових комунікацій

№ п/п	Специфіка поведінки цільових клієнтів	Канали комунікацій, якими користуються цільові клієнти	Ключові позиції, обрані для позиціонування	Завдання рекламного повідомлення	Концепція рекламного звернення
1	Цільові клієнти – торгові мережі, що займаються продажем товарів під споживчий кредит. Вони цікавляться новинками на ринку кредитних послуг.	Форуми, пряма комунікація	Позиція надання якісних послуг за меншу плату ніж на ринку	Повідомити потенційних клієнтів про новий продукт на ринку, сформувати імідж, збільшити довіру до продукту	Збільшуємо кількість продаж за рахунок продажу товарів під кредит

Висновок до розділу

В розділі зроблено аналіз побудови та виведення на ринок стартап-проекту програмного забезпечення, що було розроблено під час роботи над магістерською дисертацією.

В ході аналізу, було виявлено, що проект має можливість ринкової комерціалізації. Прогнозується високий попит на продукт та висока рентабельність.

Було проаналізовано ринок, виявлено конкурентів та побудовано стратегію виходу на конкурентний ринок.

ВИСНОВКИ

Робота присвячена побудові прогнозуючих моделей для прогнозування відповідей банків.

Проведено огляд предметної області. Після огляду поняття скорингу, історії розвитку скорингу, обмежень, які діють на скоринг та перспективи розвитку в Україні, було проведено аналіз і розглянуто існуючі рішення. Порівняно рішення задач, виділено недоліки існуючих рішень та описано результати огляду рішень. Проведено огляд існуючих аналогів систем, які є лідерами на ринку аналізу даних в банківській сфері, такі як SAS Enterprise, IBM SPSS, Deductor. Сформульовано вимоги до системи, що розробляється. Наведено опис бізнес-процесів, сформульовано акторів та функції системи.

Розглянуто існуючі методи прогнозування, розглянуто методи лінійно ймовірнісної моделі, такі як логістична регресія та скорингова карта. Побудовано скорингові моделі для оцінки шансів отримання позитивної відповіді від банку. Наведено приклади застосування моделей. Розглянуто способи оцінки скорингових моделей. Проаналізовано отримані результати та показано, що розроблена схема вирішення задачі може застосовуватись в системах для підбору кредитних пропозицій.

Розглянуто реалізацію інформаційно-пошукової системи. Представлено схеми, по яким застосунок буде передавати дані банку, та отримувати від них відповіді. Описано загальну схему роботи програмного забезпечення.

Розроблено інформаційно-пошукову систему з використанням .NET Framework у середовищі JetBrains Rider та використано базу даних SQL Server. В системі реалізовано скорингову модель для прогнозування відповідей банків. Задано загальні вимоги до технічного забезпечення.

Зроблено аналіз побудови та виведення на ринок стартап-проекту програмного забезпечення, що було розроблено під час роботи над

магістерською дисертацією. В ході аналізу, виявлено, що проєкт має можливість ринкової комерціалізації. Прогнозується високий попит на продукт та висока рентабельність. Проаналізовано ринок, виявлено конкурентів та побудовано стратегію виходу на конкурентний ринок.

Розроблений програмний продукт показав прийнятні результати, що підтверджує раціональність використання обраного методу.

ПЕРЕЛІК ПОСИЛАНЬ

1. Зверев О. А. Система продаж банківських продуктів як невід’ємна елемент ринкового механізму в банківській сфері / О. А. Зверев. // Фінанси і кредит. – 2004. – №14. – С. 3–9.
2. Рикова І. Н. Ринок нових кредитних продуктів: проблеми і перспективи / І. Н. Рикова. // Фінанси і кредит. – 2007. – №32. – С. 11–22.
3. Банних А. А. Методика оцінки якості скорингових моделей з урахуванням доходності кредиту / А. А. Банних. // Економіка і право. – 2014. – №3. – С. 21–24.
4. Андрєєва Г. А. Скоринг як метод оцінки кредитного ризику [Електронний ресурс] / Г. А. Андрєєва // Банківські технології. – 2018. – Режим доступу до ресурсу:
<https://www.cfin.ru/finanalysis/banks/scoring.shtml>.
5. Кожухівська О.А. Прогнозування ризиків кредитування фізичних осіб за математичними моделями / О. А. Кожухівська // Наукобачення. – 2014. – №2. – С. 60–82.
6. Шмелева А. Г. Програмна модель оцінки кредитоспроможності клієнтів з використанням алгоритмів штучного інтелекту / А. Г. Шмелева. // Праці НГТУ ім. Алексєєва. – 2018. – №3. – С. 72–79.
7. Сорокін А. С. Побудова скорингових моделей з використанням логістичної регресії / А. С. Сорокін. // Наукобачення. – 2014. – №2. – С. 1–28.
8. Маслова Л. А. Кредитний скоринг. Методи побудови скорингових моделей / Л. А. Маслова. // Наукобачення. – 2014. – №2. – С. 98–104.

9. Самойлова С. С. Скорингові моделі оцінки кредитного ризику / С. С. Самойлова, М. А. Курочка. // Соціально економічні явища і процеси. – 2014. – С. 1–3.
10. Кас'янчук А. С. Інформаційно-пошукова система для підбору кредитних пропозицій клієнтам торговельних мереж / А. С. Кас'янчук. // V Всеукраїнська науково-практична конференція молодих вчених та студентів «Інформаційні системи та технології управління – ІСТУ-2020» – 2020. – №1. – С. 81–85.
11. Кас'янчук А. С. Використання інформаційних технологій в кредитуванні клієнтів торгових мереж / А. С. Кас'янчук. // IX Международная научно-практическая конференция “PERSPECTIVES OF WORLD SCIENCE AND EDUCATION”. – 2020. – №2. – С. 575–579.
12. Лагарєв Д. Г. Особливості побудови скорингової моделі на основі аналітичної платформи Deductor / Д. Г. Лагарєв, І. В. Бондарєва. // Науково-технічний вісник Брянського державного державного університету, 2017. – 2017. – №1. – С. 81–85.
13. Metanit. ASP.NET WebAPI [Електронний ресурс]. – Режим доступу: https://metanit.com/sharp/aspnet_webapi/1.1.php - Дата доступу: 20.02.2020.
14. React Tutorial [Електронний ресурс]. – Режим доступу: <https://vuejs.org/v2/guide/> - Дата доступу: 08.02.2020.
15. Microsoft Docs. SignalR – Режим доступу: <https://docs.microsoft.com/ru-ru/aspnet/core/tutorials/signalr?view=aspnetcore-2.2&tabs=visual-studio> - Дата доступу: 20.01.2020.
16. Professor Web [Електронний ресурс]. – Режим доступу: <http://professorweb.ru> - Дата доступу: 22.03.2020.

17. Разработка более быстрых приложений на Vue.js [Электронный ресурс]. – Режим доступа
<https://habr.com/ru/company/ruvds/blog/487684/> - Дата доступа:
22.05.2020.
18. VuetifyJs manual [Электронный ресурс]. – Режим доступа:
<https://vuetifyjs.com/en/> - Дата доступа: 12.03.2020.
19. SAS Enterprise Miner [Электронный ресурс]. – Режим доступа:
<https://korusconsulting.ru/platforms/advanced-analytics/sas/sas-enterprise-miner/> - Дата доступа: 12.11.2020.
20. Магнус Я.Р., Катишев П.К., Пересецкого А.А. Эконометрика. Початковий курс, 2007. - 55 с.

ДОДАТОК А

Графічний матеріал

Схема структурна бази даних 1

Схема структурна бази даних 2

Схема структурна бази даних 3

Схема роботи системи

Схема структурна варіантів використання

Схема роботи Middle-ware системи

Схема роботи моніторингу сервісів банків